



PHD

## Major Evolutionary Trends

Hughes, Martin

*Award date:*  
2013

*Awarding institution:*  
University of Bath

[Link to publication](#)

## Alternative formats

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

Copyright of this thesis rests with the author. Access is subject to the above licence, if given. If no licence is specified above, original content in this thesis is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC-ND 4.0) Licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>). Any third-party copyright material present remains the property of its respective owner(s) and is licensed under its existing terms.

### Take down policy

If you consider content within Bath's Research Portal to be in breach of UK law, please contact: [openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk) with the details. Your claim will be investigated and, where appropriate, the item will be removed from public view as soon as possible.



# ***Major Evolutionary Trends***

Martin Hughes

A thesis submitted for the degree of Doctor of

Philosophy

University of Bath

Department of Biology and Biochemistry

October 2013

## **COPYRIGHT**

Attention is drawn to the fact that copyright of this thesis rests with the author. A copy of this thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that they must not copy it or use material from it except as permitted by law or with the consent of the author.

*This thesis may be made available for consultation  
within the University Library and may be photocopied  
or lent to other libraries for the purposes of consultation  
with effect from.....*

*Signed on behalf of the Faculty of Science, Department of Biology and Biochemistry*

# ***Acknowledgements***

I would like to thank my supervisor Dr Matthew Wills of the University of Bath and colleague Dr Sylvain Gerber of the University of Cambridge first and foremost. During my PhD research, they have been there for me, ready to provide me with all of the advice I ever needed when carrying out the analyses making up this thesis. I also thank them for putting up with the seemingly endless questions I had during those four years. I would also like to thank Dr Joseph Carter of the University of North Carolina in contributing data towards the bivalve chapters of this thesis.

I also thank the researchers in the Biodiversity Lab of the University of Bath and more generally my peers in the University of Bath Department of Biology and Biochemistry. All have made me feel welcome during my study. They helped me when I was stumped using R, needed to practice a talk or just needed to brainstorm an idea. Some of my peers were not able to see my four years through with me as they moved on to bigger and better positions, but while they were part of the department they made my time there that much more academically stimulating and fun. To give them their equal credit I acknowledge them in alphabetical order: Anne O'Connor, Anne Gesell, Dr Araceli Argüelles Ticó, Ross Mounce, Dr Fiona Burns, Dr James St Clair, Dr John Burnside, Dr Kate Ashbrook, Dr Katie Davies, Jorge Parra, Dr Natalie dos Remedios, Dr Peter Long, Dr Rene van Dijk and Dr Robert Sansom, and Weihao Zhong. I also acknowledge Gaby Miron whose life was tragically cut short. Her advice given to me when we shared demonstrating duties was invaluable.

I would like to thank the people in control of the Bath Aquila server for their help in using the service to conduct some of my analyses. The Aquila server allowed me to process my data in a large and semi automated fashion. I would like to thank Jessica Jones especially for her role in advising me in the correct code to make use of the server. I thank the Leverhulme Trust as the sole financial funder of the first 3 years of my study at the University of Bath. I thank Mirus Wales in providing part-time work during my 4th year so that I could carry on to finishing my thesis. Finally I would like to acknowledge my family Joy Hughes, Michael Hughes, Clare Hughes and my girlfriend Gemma Jenkins for all their love and support during the 4th year of my study and writing up the thesis. I would also like to thank my two nieces Chloe and Maisie Hughes for teaching me that play time is as important as work.

## ***Statement of Collaboration***

The following statements define the contributions made by those involved in the construction of this thesis and all papers published prior to its completion in which Martin Hughes contributed as an author. The primary R script of this thesis used in the editing of raw data matrices, building distance and principal coordinate matrices, analysing disparity patterns and graphing the results was the product of a collaborative effort between Martin Hughes and Dr Sylvain Gerber (University of Cambridge) while he was a post doctoral researcher of the University of Bath. Further modifications of the R script to take into account data downloaded from the Paleobiology Database, analysing disparity patterns across latitude and graphing those results were made by Martin Hughes. Some R functions were produced by Dr Graeme Lloyd (University of Oxford) and Dr John Alroy (Macquarie University) and are referenced appropriately both in the main text of this thesis and in the notes within the R script. Modifications to these functions were made by Martin Hughes. Modifications of the primary R script to allow analyses to be processed via the Bath Aquila server were made by Martin Hughes. The analyses themselves were all done by Martin Hughes. The graphs shown in Figure 3.3A&B (Chapter 3) were produced by Dr Sylvain Gerber. The image of a eurypterid in Figure 1.4 (Chapter 1) was made by Dr Matthew Wills. All homoplasy excess ratio results were also produced by Dr Matthew Wills.

The data of Chapter 1 were mined from the scientific literature, and requested from primary authors by Martin Hughes. The stratigraphic range data for Bivalvia that originates from an *in prep* revision of the Treatise of Invertebrate Palaeontology, as well as a discrete character morphological data matrix of Palaeozoic Bivalvia was provided by Dr Joseph Carter of the University of North Carolina. All analyses that used that data were conducted by Martin Hughes. Additional stratigraphic range data were obtained by Martin Hughes using the primary literature and online resources which are cited accordingly in the main text of Chapters 2-4. The writing of the PNAS paper (Hughes *et al.* 2013) that makes up part of Chapter 1 was a collaborative effort between Martin Hughes, Dr Matthew Wills and Dr Sylvain Gerber. All parts of Chapter 1 that do not appear in the PNAS paper were written by Martin Hughes. The writing of Chapters 2-4 and the rest of the thesis was done by Martin Hughes.



# List of Figures

## **Chapter 1:**

- Figure 1.0 Counts of the number of disparity studies per year.
- Figure 1.1 Influence of disparity papers through time.
- Figure 1.2 Counts of disparity studies showing the main biases in the literature.
- Figure 1.3 Frequency of data types used in disparity literature.
- Figure 1.4 Calculation of the disparity profile of clades.
- Figure 1.5 Number of problematic characters vs. total number of characters per dataset.
- Figure 1.6 Theoretical disparity profiles.
- Figure 1.7 Summary of the process of ancillary test on disparity profiles.
- Figure 1.8 Counts of disparity by Era for all 98 extinct clades (CG above or below 0.5).
- Figure 1.9 Counts of disparity by Era for all 98 extinct clades (CG split into 4 main categories).
- Figure 1.10 Centre of gravity ( $CG_{scaled}$ ) values for all 98 datasets across the Phanerozoic.
- Figure 1.11 Scatter plot of clade rank against time of clade origination.
- Figure 1.12 Selection of correlations between CG and taxonomic rank, number of OTUs or clade age.
- Figure 1.13 Ordered dataset CG plotted through time.
- Figure 1.14 of datasets that have used unordered vs. ordered characters states.
- Figure 1.15 of the effect that ordered vs. unordered character data has on the individual disparity profiles.
- Figure 1.16 Counts of CG in each Era (CG above or below 0.5), before and after mass extinction clades removed.
- Figure 1.17 Counts of disparity CG category by Era (CG split into 4 main categories), before and after mass extinction clades removed.
- Figure 1.18 Groups terminating at one of the 'big five' mass extinction events vs. those extinct at other times or extant (CG above or below 0.5).
- Figure 1.19 Clades terminating at one of the 'big five' mass extinction events vs. those extinct at other times (CG split into 4 main categories).
- Figure 1.20 Temporal distances (number of stages) of the  $CG_{scaled}$  from mass extinction events.
- Figure 1.21 Disparity CG categories for those clades evolving in the wake of mass extinctions.
- Figure 1.22 Distribution of vertebrate vs. Invertebrate groups ( $CG_{scaled}$ ) across time.
- Figure 1.23 Counts of CG profiles of all 98 extinct clades split by invertebrate and vertebrate group (CG split into 4 main categories).
- Figure 1.24 Distribution of data matrix 'shape'.

## **Chapter 2:**

Figure 2.0 Overlap and/or difference in stratigraphic range.

Figure 2.1 Diversity estimates (raw counts of genera in each bin) for Bivalvia across the Phanerozoic (Treatise, Sepkoski and Paleobiology Database).

Figure 2.2 Diversity estimates (raw counts of genera in each bin) for Bivalvia (Paleobiology Database).

Figure 2.3 Diversity estimates (raw counts of genera in each bin) for Bivalves across the Phanerozoic calculated from the Treatise.

Figure 2.4 Diversity estimates (raw counts of genera in each bin) for Bivalves across the Phanerozoic calculated from the Treatise (modified).

Figure 2.5 Diversity estimates (number of genera) for Bivalvia across the Phanerozoic as calculated from the Paleobiology Database (Shareholder quorum subsampling).

Figure 2.6 Diversity estimates (number of genera) for Protobranchia across the Phanerozoic (Shareholder quorum subsampling).

Figure 2.7 Diversity estimates (number of genera) for Heteroconchia across the Phanerozoic (Shareholder quorum subsampling).

Figure 2.8 Diversity estimates (number of genera) for Pteriomorphia across the Phanerozoic (Shareholder quorum subsampling).

Figure 2.9 Diversity of bivalves using the Treatise ranges and the range uncertainty method.

Figure 2.10 Number of formations and number of collections through the Phanerozoic.

Figure 2.11 Generic diversity (traditional counts) vs. the diversity expected given the number of collections through the Phanerozoic.

Figure 2.12 Generic diversity (traditional counts) vs. the diversity expected given the number of formations through the Phanerozoic.

Figure 2.13 Detrended Phanerozoic generic diversity (traditional count) using numbers of collections as a sampling proxy.

Figure 2.14 Detrended Phanerozoic generic diversity (traditional counts) using numbers of formations as a sampling proxy.

Figure 2.15 Counts of the life habits occupied by bivalves across the Phanerozoic as recorded in the Paleobiology Database.

Figure 2.16 Counts of the assumed diet of bivalves across the Phanerozoic as recorded in the Paleobiology Database.

Figure 2.17 Changes in sea level across the Phanerozoic.

### **Chapter 3:**

Figure 3.0 Bivalve anatomy.

Figure 3.1 Range of hinge dentition and ligament shape.

Figure 3.2 Range of life habits of marine bivalves.

Figure 3.3 PCO distances vs. original distances.

Figure 3.4 Choosing numbers of PCO axes for clustering analysis of Palaeozoic bivalves.

Figure 3.5 Choosing numbers of PCO axes for clustering analysis for Protobranchia.

Figure 3.6 Choosing numbers of PCO axes for clustering analysis for Heteroconchia.

Figure 3.7 Choosing numbers of PCO axes for clustering analysis for Pteriomorphia.

Figure 3.8 Estimates (traditional counts) for Bivalvia, Protobranchia, Heteroconchia and Pteriomorphia (Mixed range).

Figure 3.9 Estimates (traditional counts) for Bivalvia, Protobranchia, Heteroconchia and Pteriomorphia (Treatise range).

Figure 3.10 Comparison of generic diversity (traditional count) of bivalves across the Palaeozoic.

Figure 3.11 Diversity (SQS) of the Bivalvia and constituent subclades.

Figure 3.12 Detrended diversity of bivalves through the Palaeozoic.

Figure 3.13 Occupation of bivalves in morphospace along the first three principal coordinates.

Figure 3.14 Morphospace occupations of bivalves along the first three principal coordinates across selected stages (Treatise range).

Figure 3.15 Morphospace occupations of bivalves along the first three principal coordinates across selected stages (Mixed range).

Figure 3.16 Morphospace occupations (convex hull) of bivalve genera across the Palaeozoic (Mixed Range).

Figure 3.17 Morphospace occupation and disparity in post-Palaeozoic bivalves (End Permian vs. Recent).

Figure 3.18 Effect of character choice on morphospace occupation.

Figure 3.19 Morphospace occupations (convex hull) of Bivalvia across the selected parts of the Palaeozoic (96 character matrix).

Figure 3.20 Intensity of clustering of subclades across the Palaeozoic based on nearest neighbour analysis of the first 20 PCO axes.

Figure 3.21 Clustering Intensity across the Palaeozoic using nearest neighbour analysis. The numbers of PCO axes are: 10, 15 and 20 axes.

Figure 3.22 Clustering Intensity across the Palaeozoic using nearest neighbour analysis. The number of PCO axes for each subclade reflects the number of axes calculable using uncorrected PCO.

Figure 3.23 Principal points analysis showing the number of principal points needed to fit the distribution of genera within the Palaeozoic bivalve morphospace.

Figure 3.24 Principal points analysis of bivalves and the three subclades using the first 20 PCO axes.

Figure 3.25 Principal points analysis of the three subclades using the first 10, 15, and 20 PCO axes respectively.

Figure 3.26 Disparity (sum of variances) trajectories for Palaeozoic bivalves using different stratigraphic binning regimes.

Figure 3.27 Disparity (sum of variances) trajectories for Palaeozoic Protobranchia using different stratigraphic binning regimes.

Figure 3.28 Disparity (sum of variances) trajectories for Palaeozoic Heteroconchia using different stratigraphic binning regimes.

Figure 3.29 Disparity (sum of variances) trajectories for Palaeozoic Pteriomorphia using different stratigraphic binning regimes.

Figure 3.30 Partial disparities of subclades across the Palaeozoic (Mixed Range).

Figure 3.31 Partial disparities of subclades across the Palaeozoic (Treatise Range).

#### ***Chapter 4:***

Figure 4.0 Latitudinal diversity gradient for the marine biota across the Phanerozoic.

Figure 4.1 Mollweide globes showing the position of the palaeocontinents in Palaeozoic.

Figure 4.2 Latitudinal diversity gradient of modern day Bivalvia.

Figure 4.3 Hypothetical examples showing the gap regime for the palaeolatitudinal data downloaded from the Paleobiology Database.

Figure 4.4 Three dimensional plot showing the change in disparity and diversity with time and latitude.

Figure 4.5 Disparity profiles for bivalves across palaeolatitude at specific stages of the Palaeozoic.

Figure 4.6 Detrended diversity across latitude within each period.

Figure 4.7 Palaeozoic bivalve disparities vs. latitude (Mixed Range).

Figure 4.8 Morphospace occupations (convex hull) of genera included in the latitudinal subgroups across specific Stages of the Palaeozoic.

Figure 4.9 Partial disparities based on latitudinal subgroup across the Palaeozoic.

## List of Tables

### **Chapter 3:**

Table.3.0 Genera and characters retained in each data treatment (Mixed Range).

Table.3.1 Genera and characters retained in each data treatment (Treatise).

Table.3.2 Table showing the values given to morphological characters originally coded as letters in the original matrix of Carter *et al.* (2006).

Table.3.3 Marginal likelihood ratio between disparity values of successive Stages.

Table.3.4 Marginal likelihood ratio between disparity values of successive Epochs.

Table.3.5 Marginal likelihood ratio between disparity values of successive Periods.

Table.3.6 Marginal likelihood ratio between disparity values across the boundaries of the End Ordovician and Late Devonian mass extinctions events.

Table.3.7 Results of Spearman Rank test between disparity and time.

Table.3.8 Fit of each evolutionary model to the data (bivalves and each subclade) based on likelihood analysis.

Table.3.9 Fit of each evolutionary model to the data (bivalves and Protobranchia) based on likelihood analysis.

Table.3.10 Fit of each evolutionary model to the data (Heteroconchia and Pteriomorphia) based on likelihood analysis.

## List of Abbreviations

GRW Generalised random walk

OTUs Operational taxonomic units

OU Ornstein-Uhlenbeck process

PBDB Paleobiology Database

SQS Shareholder quorum subsampling

STA Stasis

URW Unbiased random walk

# Table of Contents

Cover Page.....	1
Acknowledgments.....	2
Statement of Collaboration.....	3
List of Figures.....	4
List of Tables.....	8
List of Abbreviations.....	8
Table of Contents.....	9
Summary.....	10
Thesis Introduction & Chapter Breakdown.....	11
Chapter 1 <i>Disparity Across the Phanerozoic</i> .....	13
1.1 Introduction.....	14
1.4 Materials & Methods.....	24
1.14 Results.....	36
1.15 Discussion.....	49
1.16 Conclusions & Future Prospects.....	51
Chapter 2 <i>The Diversity of Bivalves Across the Phanerozoic</i> .....	52
2.1 Introduction.....	53
2.4 Materials & Methods.....	58
2.6 Results.....	63
2.16 Discussion.....	89
2.17 Conclusions & Future Prospects.....	96
Chapter 3 <i>The Palaeozoic Disparity of Bivalves</i> .....	98
3.1 Introduction.....	99
3.4 Materials & Methods.....	106
3.15 Results.....	126
3.28 Discussion.....	171
3.29 Conclusions & Future Prospects.....	175
Chapter 4 <i>The Disparity of Bivalves Across the Globe</i> .....	178
4.1 Introduction.....	179
4.2 Materials & Methods.....	183
4.6 Results.....	187
4.10 Discussion.....	198
4.11 Conclusion & Future Prospects.....	200
Conclusions & Future Work.....	202
References.....	210
Appendices.....	230

# ***Summary***

Palaeontological data are essential for determining patterns of biological diversity through geological time, enabling the investigation of important macroevolutionary events such as mass extinctions and explosive radiations. Most studies utilise proxies of taxonomic diversity. A more complex undertaking is to assess patterns of morphological variety (disparity) through time, revealing the manner in which groups evolved through their 'design space'. Many published studies indicate clades tend to reach their maximum disparity early in their evolutionary history. Whether this is a real biological pattern has yet to be tested. Chapter 1 tackles the evolution of disparity in metazoans across the Phanerozoic. The results of a meta-analysis of disparity in 98 extinct clades indicate early high disparity is the most prevalent pattern across the Phanerozoic but finds no clear trends through the Phanerozoic. Mass extinction ended clades were the exception, tending to result in late high disparity.

Chapters 2-4 focus on the clade Bivalvia for disparity and diversity analysis. Bivalves are ecologically and taxonomically diverse and have an excellent fossil record but have not been scrutinised using the latest diversity techniques, and have been untouched by disparity analysis. Chapter 2 uses the most up to date stratigraphic ranges and techniques to revise the bivalve Phanerozoic diversity curve. The results show bivalve Phanerozoic diversity is robust to the sampling and fossil record biases examined. Chapter 3 uses data provided as part of collaboration between Martin Hughes, Dr Joseph Carter (University of North Carolina) and Dr Matthew Wills (University of Bath) to address the disparity of bivalves across the Palaeozoic. The results find disparity rises across time but not decreased by mass extinctions. Chapter 4 conducts the first large scale analysis of disparity across latitude. The results find that bivalve disparity across latitude is unchanging and stable compared to the steep gradient of bivalve diversity.

# ***Thesis Introduction & Chapter Breakdown***

The fossil record provides the only way to illuminate many major trends in biological evolution above the level of species (macroevolution). Such trends include changes in taxonomic diversity, morphological diversity (disparity), complexity or body size over geological timescales and/or distance from the tropics (Sepkoski 1984; Gould 1989; Valentine *et al.* 1994; Foote 1993b; Clauset & Erwin 2008; Krug *et al.* 2009). Studying these large scale patterns is enticing as the results offer insight into how evolutionary processes have affected biodiversity over the last 542 million years. By identifying and analysing the patterns that operate at the macroevolutionary level one can begin to recognise the evolutionary rules that govern those processes (McShea 1998; McShea & Brandon 2010). The focus of this thesis is twofold. The thesis first focuses on the macroevolutionary patterns of disparity, the morphological differentiation of taxa, across the Metazoa through the Phanerozoic. The thesis then focuses on a specific clade to identify the trends in diversity and disparity specific to particular time intervals and latitudinal bins.

A number of studies have suggested that many clades reach their highest level of disparity early in their evolutionary history (Ruta *et al.* 2006; Wills *et al.* 1994; Wills & Fortey 2000; Erwin 2007; Foote 1997; Gould 1989; Wagner 1997; Foote 1993b; Valentine *et al.* 1996), an assumption born from a slow build up of disparity related literature. Whether this assumption is a real biological pattern has not been rigorously tested. The first chapter tackles the assumption that early high disparity is a general rule across all clades; and a pattern most prevalent in the Palaeozoic (Gould 1989; Valentine 1995; Wills & Fortey 2000; Wagner 2010). The problem with the current consensus is the lack of a rigorous meta-analytical framework from which to base this conclusion. Current understanding of the evolution of disparity through time is born from multiple results stemming from different studies using various kinds of data, metrics and methods. Pooling these studies without any consideration for their compatibility with one another is problematic. There is the bias in the kinds of clades represented by these studies, with greater influence given to older, highly cited literature based on a selection of Palaeozoic originating invertebrate clades (Gould 1989; Briggs *et al.* 1992; Foote 1993b). There is also no clear definition of the term “early” in regard to a clades evolutionary history.

To test whether early high disparity is truly an evolutionary rule, Chapter 1 analyses discrete character datasets mined from the published literature, calculating disparity (sum of variances) across time under a meta-analytical framework. Disparity patterns are categorised using the centre of gravity metric, providing a temporal measure of “early” and “late” highest disparity along a clades evolutionary history. As disparity patterns were tracked across the whole of the Phanerozoic, Chapter 1 also explores the



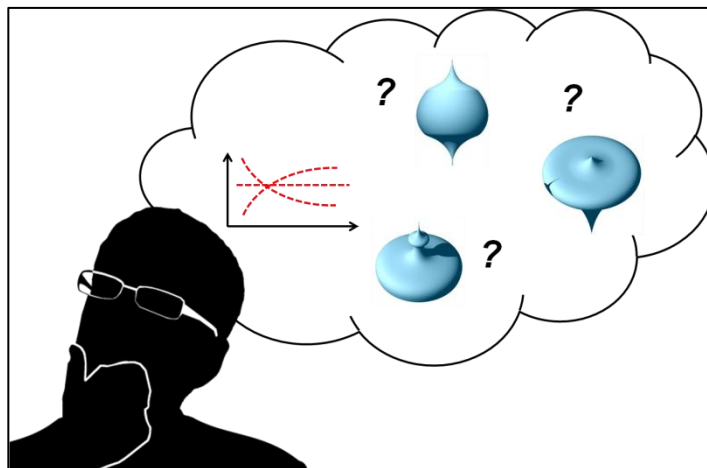
influence that the 5 biggest mass extinction events had on shaping disparity patterns in clades terminating at those events or originating in their wake.

The next three chapters focus on a specific clade: the Bivalvia. They are an excellent model clade for macroecological and macroevolutionary studies of diversity (Krug *et al.* 2009; Jablonski *et al.* 2013); but have not received much attention as a model clade for disparity (Grey *et al.* 2010; Kolbe *et al.* 2011; Anderson *et al.* 2010; Schneider *et al.* 2010). Bivalves have much to offer to the study of macroevolutionary history. They are a highly diverse modern clade with an evolutionary history and excellent fossil record that stretches back to the beginning of the Ordovician. They are survivors of every mass extinction to date and have explored an array of aquatic habitats and ecological roles (Ponder & Lindberg 2008). Despite their mostly sessile nature, they are found globally, and able to profoundly influence their environment (Sousa *et al.* 2009). They have a body form unique amongst the molluscs, a soft body encased within two calcified valves. Their morphological evolution is contradictory as they have both conserved the same basic body form since their origination in the Cambrian and evolved a wide range of morphological structures within this conserved framework.

By tracking how bivalves have diversified morphologically over geological time Chapters 2-4 provide the empirical evidence which can assist in the continuing effort to understand the mechanics of biological evolution. Chapter 2 assesses the robustness of the bivalve Phanerozoic diversity curve to revisions in stratigraphic data using three sources: the Sepkoski Database (Sepkoski 2002), the Paleobiology Database (Alroy *et al.* 2001), and a revised edition of the Treatise of Invertebrate Palaeontology (*in prep*). This is important as the current understanding of bivalve evolutionary history stems from published data that is now well over 40 years old (Cox *et al.* 1971). The chapter tests whether the diversity curve is the result of geological, sampling or methodological artefacts by comparing three different methods of diversity estimation: traditional counting, resampling, and modelling techniques. From this new base level of understanding of bivalve diversity through the Phanerozoic, Chapter 3 then quantifies the changes in diversity and disparity in bivalves in the Palaeozoic combining the techniques of Chapters 1 and 2; investigating in detail how bivalves and their constituent subclades evolved over this important interval of their history. Chapter 4 investigates how disparity changes with latitude using Palaeozoic bivalves as the focal clade, testing whether disparity follows a gradient across latitude, and whether tropical genera are more morphologically disparate than extra-tropical genera.

# ***Chapter 1***

## ***Disparity Across the Phanerozoic***



## 1.1 Introduction

Evolution is usually characterised as an essentially contingent and unpredictable process (Ghiselin 2005). This makes it very difficult to identify general rules comparable to those that typify the other natural sciences. Nonetheless, the prospect of formulating and testing macroevolutionary generalities; i.e. evolutionary trends visible in taxonomic ranks above the level of species is extremely seductive, since they appear to offer fundamental insights into the manner in which evolutionary processes operate throughout Earth history (McShea 1998). These generalities may be built upon the complex interactions between microevolutionary processes such as genetic drift, gene flow, mutational selection between populations, and the interactions of these populations with the environment and other populations. It is difficult to fit these individual pieces together to adequately explain the patterns of macroevolution identified using fossil record, and so to understand these macroevolutionary patterns they are studied directly.

The most recognisable trend in the fossil record is the increase of diversity with time, measured using proxies of species richness (Sepkoski 1984; Benton & Emerson 2007; Alroy 2010b). Trends that run parallel to the increase in diversity through time are the increase in organismal complexity (Arthur 2006) measured by counting the number of unique cell types (Valentine *et al.* 1994) or quantifying the degree of limb specialisation (Adamowicz *et al.* 2008) through time. Other macroevolutionary rules relate to changes in body size, including Cope's rule (i.e. species tend to be larger than their ancestors) (Sander *et al.* 2011); the island rule (i.e. gigantism of small animals and dwarfism of large animals on islands); or Bergmann's rule (i.e. body size increase away from tropics) (Clauset & Erwin 2008; Lomolino *et al.* 2013). Another general rule relates to the global distribution of organisms, the so called latitudinal diversity gradient; i.e., that diversity peaks in the tropics and decreases toward the poles. These gradients have been documented hundreds of times in numerous different organisms (Hillebrand 2004) and considered one of the earth's first order biodiversity patterns (Krug *et al.* 2009).

In the book 'Biology's first law' by McShea and Brandon (2010), the authors put forward a law of biology that acts as a mechanism by which all life operates, i.e. that diversity and complexity increase with time when not the target of constraint, known as the zero force evolutionary law. They cite another putative generality as evidence for its existence. This generality, the tendency for species to reach their maximal morphological diversity (disparity) relatively early in the lifespan of their parent clade (Ruta *et al.* 2006; Wills *et al.* 1994; Wills & Fortey 2000; Erwin 2007; Foote 1997; Gould 1989; Wagner 1997; Foote 1993b; Valentine *et al.* 1996) is otherwise known as 'early high disparity'. The general pattern of disparity was born from the observation that the modern biota exhibits a wealth morphological differentiation, and has done so since the Cambrian with fossils such as *Wixwaxia* and *Anomolacaris*.

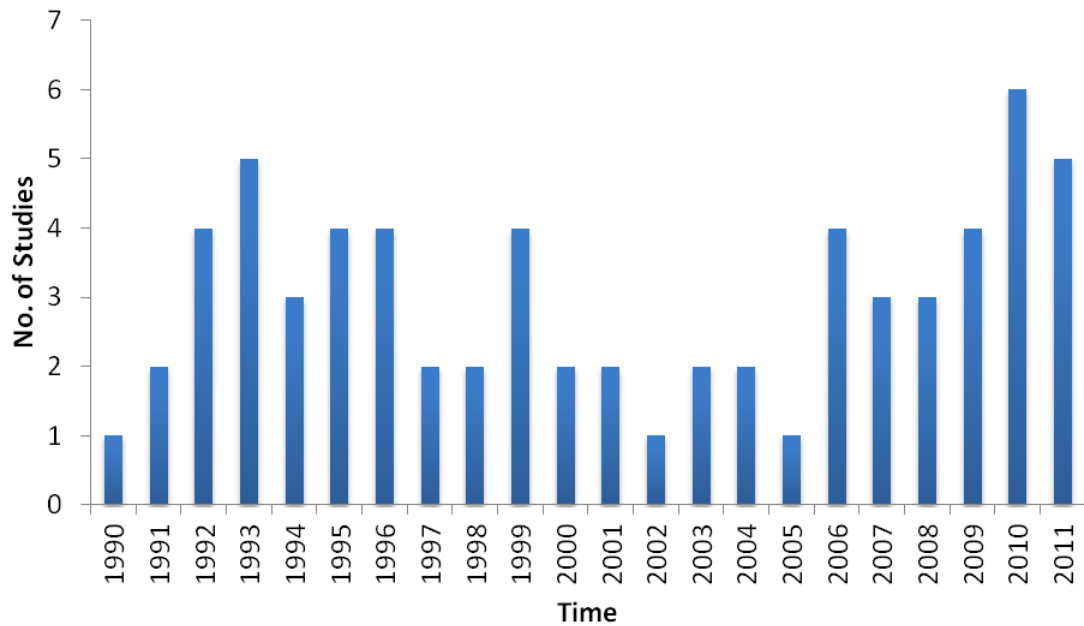


Figure.1.0 Counts of the number of disparity studies per year (1990-2011) that have looked at morphological disparity. Data compiled from Appendix.1.Table.1.

### 1.2 Disparity vs. Diversity

As an area of research, morphological disparity analysis is now over two decades old (Fig.1.0). Figure 1.0 shows counts of the number of disparity studies that have been written each year from 1990-2011. The raw data making up the plot can be found in Appendix.1.Table.1. The plot shows a surge of interest in the study of disparity in the early 1990s after the publication of a *Wonderful Life* (Gould 1989). These early papers (Gould 1989; Foote 1991a; Foote 1991b; Briggs *et al.* 1992; Gould 1993; Foote 1993b; Foote 1993a; Labandeira & Sepkoski 1993; Foote 1994a; Wills *et al.* 1994) were highly influential as seen in the number of citations since their publication as recorded on *ISI/ Web of Knowledge* or *Google Scholar* (Fig.1.1).

The primary focus of disparity analysis is to describe the change in net differences between biological forms (Wills 2001); however, the techniques involved are gradually being turned to other aspects of morphology including ecological breath (Labandeira & Eble 2007), geographic biodiversity patterns (Neige 2003; Neustupa *et al.* 2009; Grey *et al.* 2010), and allometric trajectories (Gerber *et al.* 2008; Zelditch *et al.* 2003) Although disparity analysis can use the same raw discrete character data used in some phylogenetic analyses, taxonomic diversity and morphological disparity are considered complementary but separate measures of biodiversity.. Diversity studies use morphological traits to differentiate species, in reference to some phylogenetic hypothesis, and can be used to construct diversity curves (Labandeira & Sepkoski 1993; Benton 2009).

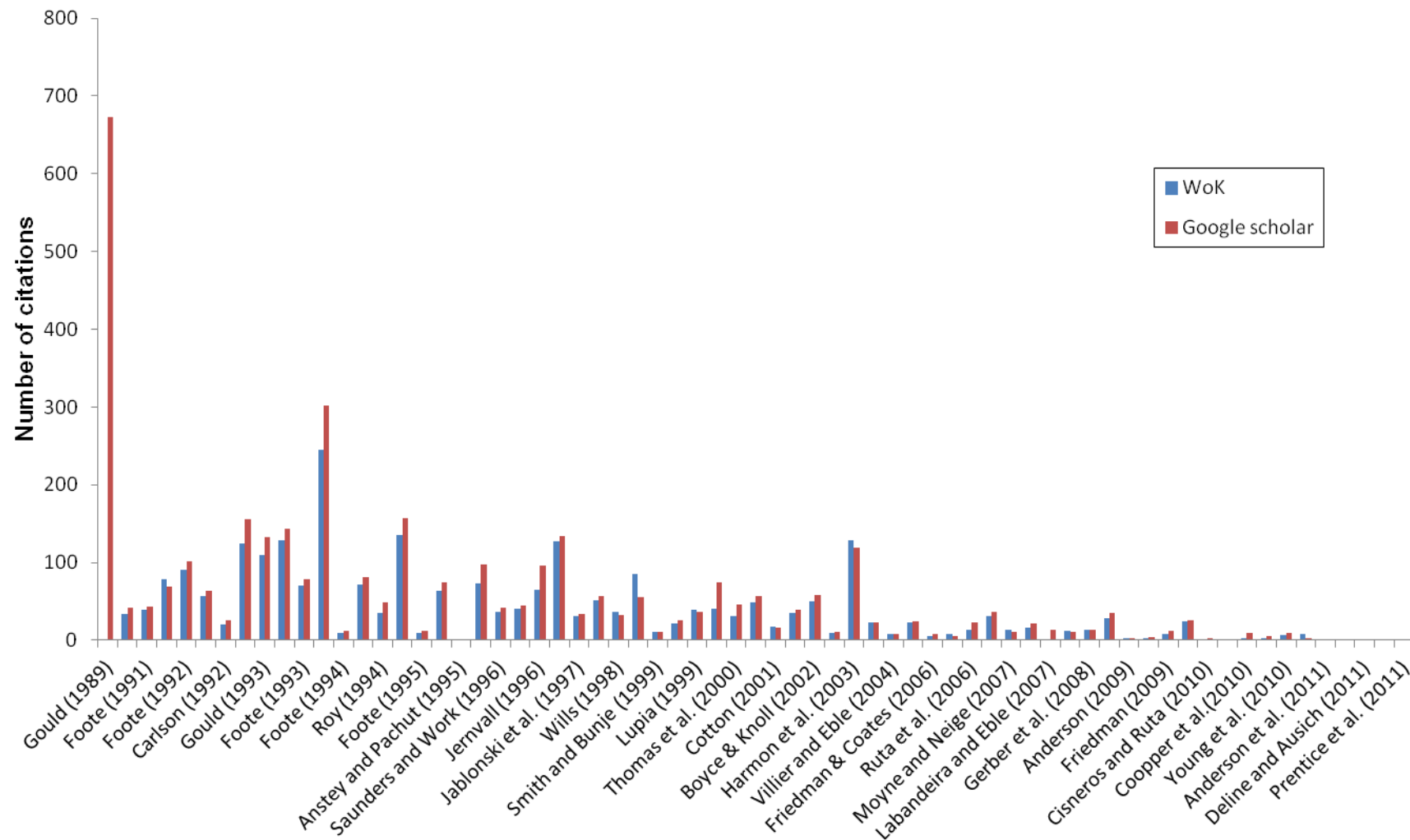


Figure.1.1 Influence of disparity papers through time based on the number of citations from ISI Web of Knowledge (blue) and Google Scholar (red). Data compiled from Appendix.1.Table.1.

Disparity is conceptually and empirically distinct from diversity. For example, a relatively small sample of species that differ greatly from one another morphologically (e.g., one species from each order of insects) is likely to be more disparate than a much larger sample of species that are morphologically more homogeneous (e.g., a thousand beetles). Among the first questions to be addressed using disparity indices was the perceived magnitude of the Cambrian ‘explosion’ (Gould 1989). From Charles Darwin onwards (Darwin 1859), evolutionary biologists have been perplexed by the apparently instantaneous first appearances of numerous phyla (a highly disparate sample of species) in the Cambrian fossil record (Wagner 2010). The subsequent discovery of previously unknown fossil groups from the Cambrian Burgess Shale and similar localities added to the enigma, prompting the radical hypothesis that the disparity of metazoans peaked in the Cambrian (Gould 1989; Gould 1991) with subsequent extinctions winnowing this down to much more modest levels soon thereafter. Authors focusing on the range of body forms originating in the Cambrian and Ediacaran found disparity to be high (Thomas *et al.* 2000; Shen *et al.* 2008).

This “inverted” cone presented the opposite of what could be considered the “traditional” view of morphological diversity through time, .i.e. a cone of increasing morphological differentiation (Wills & Fortey 2000) following in tandem with the increasing diversity through the Phanerozoic (Sepkoski 1984; Miller & Sepkoski 1988; Alroy *et al.* 2008). Surprisingly, a relatively small number of studies have tested this hypothesis directly in focal clades (Briggs *et al.* 1992; Foote 1992a; Wills 1998a; Wills *et al.* 1994; Wills & Fortey 2000). These predominantly concluded that Cambrian animal groups had a disparity comparable to that of their modern counterparts (Valentine *et al.* 1999; Thomas *et al.* 2000; Knoll & Carroll 1999; Shen *et al.* 2008). This nonetheless suggests that metazoans reached high levels of disparity relatively early in their history; the phenomenon of ‘early high disparity’.

Unfortunately, such analyses are limited for two reasons. Firstly, they discount the intervening trajectory of clade evolution. Secondly, clade history is truncated both by the present, and by the Precambrian. The truncation at the present is a problem as one cannot yet predict what future evolution will occur in extant taxa, especially over millions of years. One solution would be to consider the present as the end of the evolutionary history of extant taxa, effectively creating an artificial and final extinction. There is the potential problem that this could bias the results towards one particular pattern if the majority of clades today are still evolving and have not yet reached their highest disparity. Therefore for the analyses herein the conclusions of the results are based on extinct clades as to have some idea of their entire evolutionary history. The Precambrian is a problem as its fossil record is still enigmatic, with numerous fossil organisms (e.g. *Pteridinium*, *Charnia* and *Swartpuntia*) with no concrete connection to either each other or to the phyla appearing in the Cambrian (Waggoner & Collins 2004; Valentine *et al.* 1996).

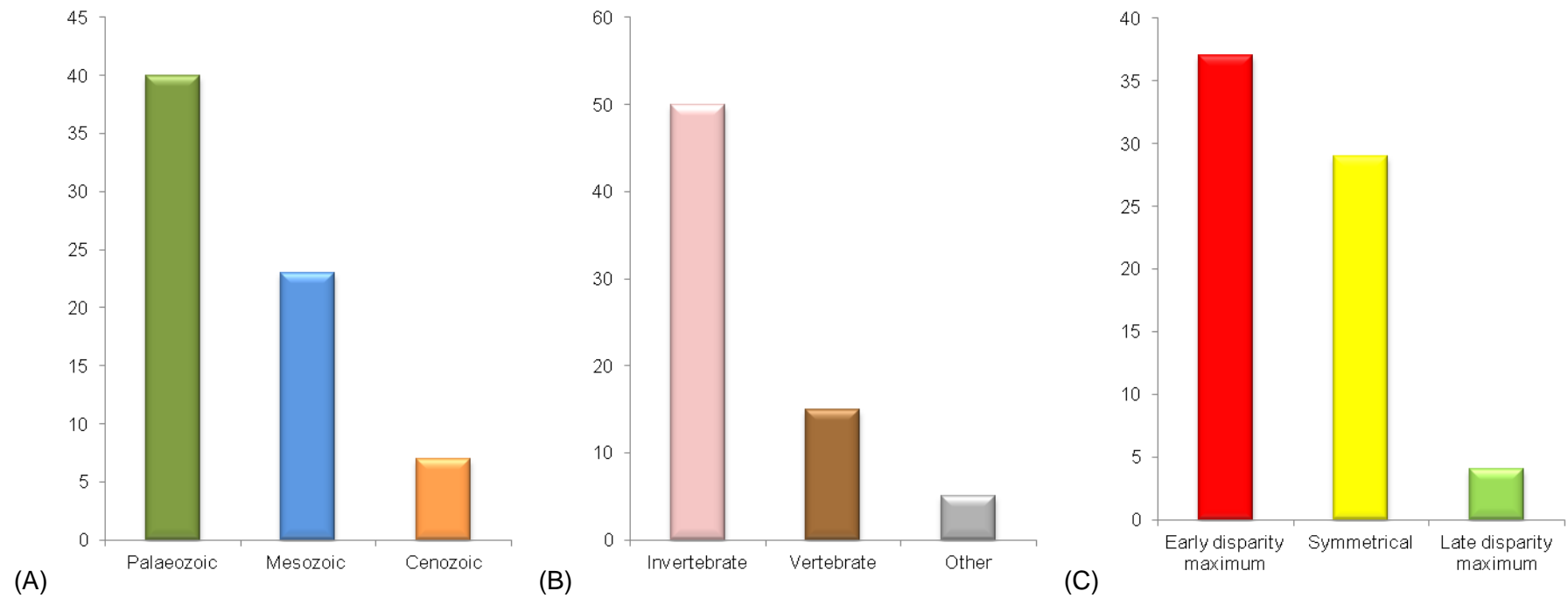


Figure.1.2 Counts of disparity studies showing the main biases in the time intervals of interest, the group of interest and the pattern of disparity as described in the source paper (Appendix.1.Table.1).

There is also the problem with the stratigraphic resolution. Some of the time intervals of the Precambrian compared to the individual time intervals in this the Phanerozoic are over 200 million years, for example the Cryogenian or Tonian (Gradstein *et al.* 2004; Ogg *et al.* 2008). Finally there is a problem with the paucity of large datasets from this period (Shen *et al.* 2008; Huntley *et al.* 2006) due to the small number of Metazoa and disagreement on the structures of those taxa of this period (Van Iken *et al.* 2013).

As a result, the focus of disparity studies has increasingly turned to clades that both originate and go extinct *within* the Phanerozoic (Gould 1991) (Fig.1.0 and Fig.1.1). Figure 1.2 shows the counts of the published studies categorised by time (the three major subdivisions of the Phanerozoic: Palaeozoic, Mesozoic and Cenozoic), group of interest (invertebrate, vertebrate and others (plants and acritarchs) and the conclusion about the pattern of disparity (“early high”, “symmetrical” or “late high”). Once again, there is a *purported tendency* for clades to evolve their most disparate forms relatively early in their histories (Wills & Fortey 2000; Erwin 2007; Foote 1997; Gould 1989; Erwin 2011; Foote 1996). The validity of this ‘early high disparity’ model has never been tested systematically. *If true*, it represents a general macroevolutionary ‘rule’ (Wagner 2010) on the broadest possible scale, and comparable to those proposed for increasing morphological complexity (Arthur 2006; Adamowicz *et al.* 2008) and increasing maximal organismal size within clades (McShea 1998; Hone & Benton 2005).

### **1.3 Disparity Literature**

Unfortunately, it is impossible to interpret published case studies meta-analytically for several reasons. Firstly, the type of data utilised is highly variable (Fig.1.3). Some of these types of data are what could be classified as “traditional” disparity data i.e. data directly describing some aspect of the shape or structure of the animal: outlines, landmarks, measurements, discrete characters, or a mixture of these. They are traditional in the sense that they take the morphologies of the study group as the raw data from which the differences between taxa can be calculated. Less conventional data that has also been used to measure disparity includes using only specific parts of the anatomy (Jernvall *et al.* 1996) or novel structures (Jablonski *et al.* 1997), counting the number of actual character states (Crepet & Niklas 2009), polymorphic characters (Webster 2007) or even using the ecological classification based on the overall morphology (insect jaws in this case) (Labandeira & Eble 2007). As Figure 1.3 shows, discrete character data has been used most often but is one of various possible kinds of data that could be collected to measure disparity.



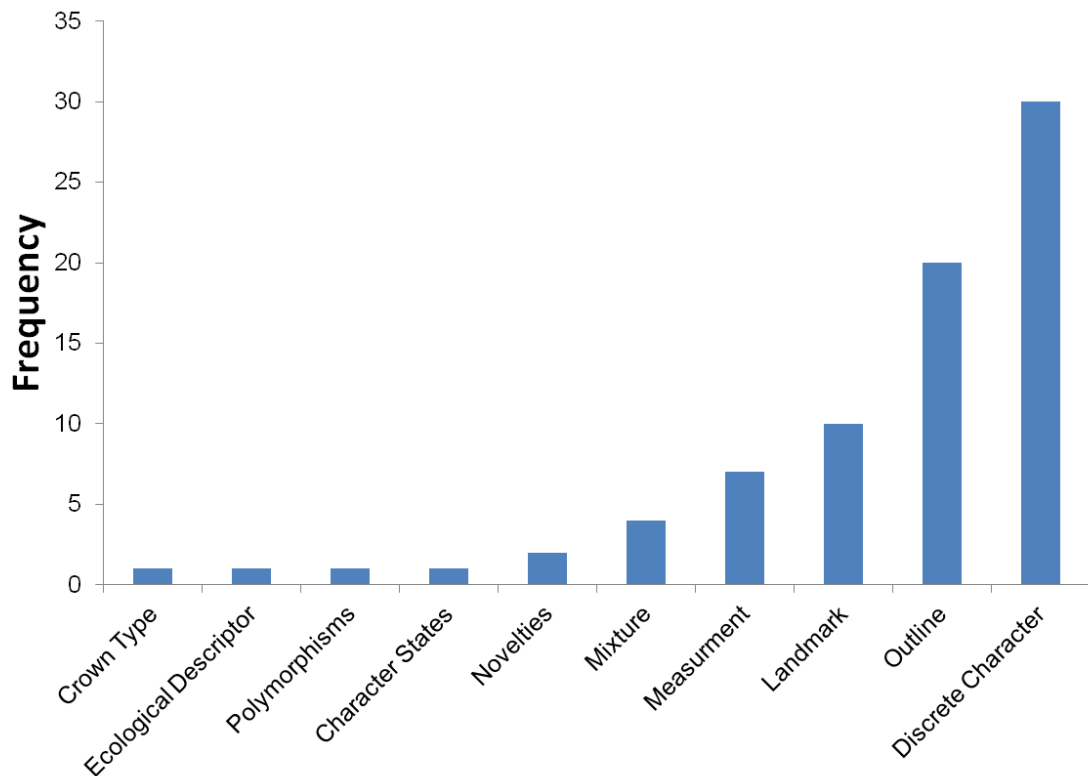


Figure.1.3: Frequency of data types used in disparity literature. (Appendix.1.Table.1)

Secondly, the manner in which these data have been visualised is equally variable, although most studies implement some form of data reduction and ordination (Wills *et al.* 1994; Erwin 2007). Specimens are typically plotted within an empirical, multidimensional space defined by morphological variables (a morphospace) (Wills 2001). There are several methods by which morphospace can be calculated and visualised: factor analysis, Fourier analysis, principal component/coordinate analysis, relative warps analysis, non metric multidimensional scaling, and canonical variate analysis. Raw measurement data can also produce a very simple morphospace.

Thirdly, there are many possible indices of morphological disparity, and these are known to describe different aspects of morphospace occupation (Ciampaglio *et al.* 2001). They fall into two broad categories: distance-base metrics, and phylogeny-based metrics, which roughly equates to the magnitude and mode based explanations of Sidlauskas (Sidlauskas 2008). Distance based metrics cannot take account of homoplasy, and by extension resolve phylogeny (Smith 1994), but are useful because they can be used to “..draw evolutionary inferences from temporal patterns in the realised dissimilarities among representatives of various lineages” (Foote 1995). Phylogeny based metrics of disparity can be used as a method for accounting for homoplasy; however, they also require the construction of a tree, which is itself only a hypothesis of phylogenetic relationship.

The most popular method of disparity analyses have been distance metrics. They can be split into two subtypes. The first types are ‘box’ like metrics which give the overall

spread and occupation of points in the morphospace. The two most popular metrics are the sum of variances and sum of ranges respectively, and their multidimensional equivalents: the product/geometric mean of variances or ranges (Foote 1991a). Variance, calculated as the sum of univariate variances, gives a measure of spread through the space; while the product/geometric mean are the variance of a user defined number of axes e.g. the first 5 principal coordinate axes. It is robust to changes in sample size and to the presence/absence of outlier forms (Ciampaglio *et al.* 2001); however, it is sensitive to uneven taxonomic practice (Wills 2001). It is equivalent to the mean squared Euclidean distance of all taxa in the space and the group centroid, or the sum of Eigen values (Eble 2000).

Another popular method used is the range, calculated as the maximum Euclidean or Manhattan distance (Ruta 2009) between specimens (Ciampaglio *et al.* 2001) indicates the total range of morphospace occupied by the most extreme forms. As a consequence it is sensitive to changes in sample size, and the presence/absence of outlier forms. Unlike variance it is immune to uneven taxonomic practice (Wills 2001). From the mid 1990s onwards more complex measures of total morphospace occupation were available, including calculating areas of convex hulls (Foote 1999; Dera *et al.* 2010) and the percentage of morphospace, i.e. the distance between partial morphospace centroids (Dommergues *et al.* 1996). These have the same kind of bias as the regular range but try to account for areas of morphospace not occupied by taxa.

The other type of distance metrics are what I define as 'net' like. The term net stems from the net like pattern of calculating distances between all points in the morphospace. These metrics calculate the distances between pairs of points or between points and some centroid. There are several subtypes. The nearest neighbour distance joins the points to their nearest neighbour in the space (Foote 1990), which can indicate clusters of points. Although superficially similar to the phylomorphospace approach discussed in Sidlauskas (Sidlauskas 2008), they differ as the former is a phenetic similarity while in the latter the pattern is based on the relationships between the taxa in a phylogeny. Others look at the mean pairwise distances between points (Wills 2001) or distances between points and some centroid (global, mean shape or out-group (Wills *et al.* 1994; Wills 1998b). These 'net' metrics are useful as they do not depend on the orientation of subgroups in relation to the overall axis of the morphospace (Wills 2001).

For a given dataset, phenetic methods can underestimate the "true" evolutionary distance between taxa by not accounting for reversals and convergences in character evolution. The earliest occurrence of phylogeny metrics of disparity was in Briggs *et al.* (Briggs *et al.* 1992). In this instance they used the patristic distance between the basal node and tip of each branch in a phylogenetic tree as a measure of disparity. The simplest method is to simply compare variation in some aspect of form between different clades. Several authors have done this: Carlson (1992) used shell properties (moment, lever arm

and force), while Jernvall (1996), and Labandeira and Sepkoski (1993) separated taxa based on classifications built around several characters of morphology (tooth and mouth parts, respectively). Disparity for these clades was then based on when the number of families which varied in these traits was greatest. Anstey and Pachut (1995) used an even simpler method, measuring the greatest disparity to be simply when the number of families was greatest. These types of approaches are good initial stepping stones for measuring disparity but must be accompanied by more quantifiable methods. Wagner (1995; 1997) used morphological separation (average patristic dissimilarity per branch) to determine the amount of morphological evolution separating species along a phylogeny, while Sidlauskas (2008) produced a morphospace of points using landmarks and calculated the distances between these points based on their phylogenetic relationships in the tree.

Wagner (1995), noted that movement of one or more landmarks in a set of measurements will be more or less significant in evolutionary terms. In his study group (gastropods) he showed the rotation of a whole set of landmarks is more likely than the movement of one landmark and more evolutionary informative. He produced transition axes, which could be used to measure the magnitude of transition in form between two species, indicating where the most evolutionary meaningful changes were occurring; the downside of the approach is it requires much prior knowledge about what is deemed to be a significant change in form.

Another range of disparity analyses have tried to link the changes in the size of animals with trees based on genetic data to assess the level of disparity within clades through time (Harmon *et al.* 2003; Burbrink & Pyron 2010; Harmon *et al.* 2010; Slater *et al.* 2010; Dornburg *et al.* 2011). Known as the morphological disparity index (MDI), the metric looks at the overall difference of disparity of a clade compared to disparity under a null model of evolution simulated through a model of Brownian motion. The null model relates to the level of character change along the branches of the tree. It was first developed by Harmon *et al.* (2003) for extant iguanian reptiles but has since been used for extant snakes and whales (Burbrink & Pyron 2010; Slater *et al.* 2010). This disparity metric plots the level of morphological evolution (in the form of limb measurements) along a molecular based tree, calculating the average pairwise Euclidean distances for the whole dataset, and then for each of the subclades in the tree. This metric is useful where gene data is available but taxonomic description has lagged behind. The metric then indicates whether the disparity along the tree is being partitioned into or between subclades. This method is currently only available for continuous data because of the method in which it simulates data (personal communication Luke J. Harmon, 2010).

Apart from those phenetic and phylogeny based distance metrics, there are also other ways to calculate disparity. A few authors have used the number of unique pairwise

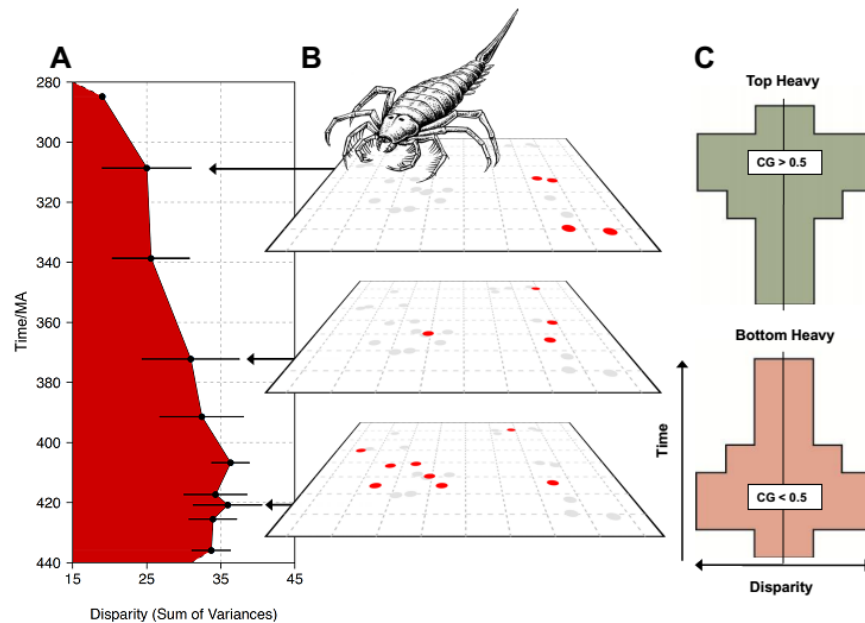


Figure 1.4. Calculating the disparity profile of clades. (A) Disparity of Stylonurina (Lamsdell *et al.* 2010) measured as the sum of variances on successive principal coordinate analyses at several time intervals. Mean of 1000 bootstrap replicates  $\pm$  standard error. (B) Distribution of taxa on the first two principal coordinates of their empirical morphospace at three of the time intervals. Red symbols indicate taxa present in the interval; gray symbols indicate taxa present in other intervals. (C) Stylised representations of significantly top-heavy (upper) and bottom-heavy (lower) asymmetrical clade disparity profiles under the idealised conditions where all time bins are of equal duration. Eurypterid courtesy of Dr Matthew A. Wills (University of Bath)

character combinations as a metric of disparity (Ciampaglio *et al.* 2001; Foote 1994a; Foote 1994b; Shen *et al.* 2008). This metric shows the number of realised character state combinations in a sample. As the number of characters increases, so does the number of possible combinations. It is also biased by sample size (Foote 1994a). In the paper by Webster (2007), the author plotted the numbers of taxa with polymorphic states across the Cambrian, arguing that those time periods with greater numbers of polymorphic states was an indication of higher morphological disparity. A problem with this method is that these polymorphic states may be an artefact of taxonomic practice and not a real signal of morphological variety.

Several of the analyses that originally spurred the debate (Wills *et al.* 1994; Briggs *et al.* 1992; Foote 1992a; Foote 1994a) utilised discrete character matrices in order to compare anatomically very disparate forms. Despite the number of other methods and metrics available, many studies have recently returned to following similar protocols as these early papers (Shen *et al.* 2008; Brusatte *et al.* 2008; Bapst *et al.* 2012; Cisneros & Ruta 2010), and I adopted these methods here as a unifying approach. Where discrete and continuous character data have been compared for the same sets of taxa (Foth *et al.* 2012), relative estimates of disparity have been similar.

## 1.4 Materials & Methods

Morphological and stratigraphic data for 98 extinct and relict clades was collected in order to answer three questions. (1) Is 'early high disparity' the dominant pattern of clade evolution across the Metazoa and throughout the Phanerozoic? (2) Is there a trend in clade disparity profile shape throughout the Phanerozoic? (3) Do clades terminating at times of mass extinction have disparity profile shapes distinguishable from clades becoming extinct at other times? I addressed all three questions using the clade centre of gravity (CG) index (Gould *et al.* 1987; Uhen 1996; Foote 1991a). This quantifies overall clade shape in a robust manner, and has previously been applied to paleontological diversity and disparity data such as in Blastozoa and Trilobita (Foote 1993b) (Appendix.1.Table.2). Values of  $CG < 0.5$  denote bottom-heavy clades, while  $CG > 0.5$  indicates top-heaviness.

Extinct clades were used as extant lineages may still be evolving, and may be at (or still approaching) their maximum disparity. Extant clades are therefore more likely to have 'flat-topped' disparity profiles, which will artefactually shift their CG upwards (Fig.1.4C) relative to that which may have pertained for the (hypothetical) entire clade history. Clades terminating at mass extinction events might be similarly truncated, and are likely to have higher CGs for similar reasons. Mass extinctions have undoubtedly influenced the manner in which clades have explored morphospaces (Jablonski 2005), but this phenomenon received little attention until recently (Bapst *et al.* 2012; Friedman 2010; Brusatte *et al.* 2012; Ausich & Deline 2012; Girard & Renaud 2012; Thorne *et al.* 2011). Moreover, only one of these studies (Brusatte *et al.* 2012) focused on extinction selectivity *per se*; all others investigated the subsequent evolution of extinction survivors. Here, I determined whether the clades going extinct coincident with one of the 'big five' mass extinction events (End Ordovician, Late Devonian (Frasnian/Famennian), End Permian, End Triassic, and End Cretaceous) had disparity profiles distinguishable from those terminating at other times.

## 1.5 Collation of Data

All research articles were identified using *ISI Web of Knowledge*, *Google Scholar*, individual journal websites and hard copy publications. Electronic copies of data matrices were obtained from supplementary materials, standard repositories or the original authors. Where electronic versions were unavailable, the data were transferred from PDF to Nexus format semi-manually and read into *PAUP\* 4.0* (Swofford 2003), from which a tab delimited text file was exported.

The initial selection criteria for data was for studies that: (1) contained phylogenies derived from discrete morphological character data; and (2) contained extinct metazoan taxa. As far as possible, the most inclusive and most recent studies that addressed the relationships of the taxa within the clade of interest were used. I picked datasets that were

not obviously derivatives or subsets of one another; containing significantly overlapping sets of characters or taxa. A total of 91 publications (Appendix.1.Table.2) contained data suitable for disparity analysis. These either (1) comprised solely extinct taxa, or (2) could be dynamically pruned in such a way (see '**1.6 Taxonomic Resolution**' below) that all remaining taxa were extinct and represented monophyla. For two datasets (Hu *et al.* 2009) I chose to use paraphyletic groups. The first was Theropoda excluding Avialae as I wished to focus on non avian dinosaurs as this group can be clearly separated due to the change in ecology of the group (Brusatte *et al.* 2012), while the second was cladid crinoids excluding other Articulata to fall in line with the classification and sampling regime followed by Foote (1999).

The 91 publications identified in this way yielded data for 98 independent clades spanning the entire Phanerozoic. Four papers (Anderson *et al.* 2011; Carlson & Fitzgerald 2007; Foote 1999; Mitchell 1987) yielded data for more than one independent clade (Appendix.1.Table.2). Of the 91 papers, just seven had previously been used to produce empirical morphospaces and the remaining 84 had never been utilised in this way. For all terminals or operational taxonomic units (OTUs) in all datasets the ages of first and last fossil occurrences were determined using individual databases, web resources and primary papers (see '**1.7 Stratigraphic Resolution**' below). Ranges were thereby coded to the stratigraphic resolution of stages, based on the 2009 International Stratigraphic chart (Adnet & Cappetta 2001; Anderson & Roopnarine 2003). Because most of the original data sets were constructed to address phylogenetic questions, the majority required some moderation prior to analysis (see below).

### **1.6 Taxonomic Resolution**

For many datasets, it was necessary to remove outgroups. This was because cladistic matrices are often constructed by consistently sampling the ingroup, but with piecemeal or subsequent additions of outgroup taxa. Outgroups were therefore removed for any of three reasons: (1) the outgroup OTUs did not constitute a recognised monophyletic clade with the ingroup (i.e., there were missing taxa between the ingroup and outgroup); (2) the outgroup OTUs were sampled at a higher taxonomic level than the ingroup; (3) less than 50% of the outgroup OTUs were sampled, albeit at the same taxonomic level as the ingroup. In practice, outgroups were usually excluded for the second and third reasons. Retaining both classes of outgroup would have the effect of comparatively undersampling both taxonomic and morphological variation in those lineages of the total clade closest to its root. Keeping outgroup taxa could also affect CG results if the outgroups were morphologically distinct from the ingroup taxa. This would increase the values of disparity in the intervals over which the outgroup ranged and possibly push the CG away from its naturally resting point for the ingroup in isolation.

In some datasets, inhomogeneity of sampling within the ingroup (e.g., some lineages were sampled at the familial level, others were sampled at the ordinal level) was also a potential problem and needed to be accounted for. It is problematic for variance-based indices of disparity (e.g., the sum of the variances on the axes of the ordination). Disproportionate sampling and splitting of some groups relative to others will result in a reduction of apparent variance, and a shift in the centroid of the empirical morphospace towards those more intensively sampled groups. Conventional taxonomic ranks only offer a first approximation to a more even sampling strategy, but where genera are coded alongside classes (for example) it is unlikely that the sampling is unbiased with respect to either morphological or taxonomic diversity.

Whenever possible the OTUs of the clade were assigned the taxonomic ranks intended by the original authors of each dataset. The next step was to assess the taxonomic ranks of those OTUs given the focus of the paper. For example, although an author may have coded fifty named species, these may actually have been selected as exemplars of fifty genera, fifty families, or fifty taxa of some higher rank. If this was not stated explicitly in the source materials, I inferred the sampling regime with reference to the author's contemporaneous papers, or (failing this) with reference to the most congruent available taxonomic scheme.

The efficiency with which this sampling scheme encoded the total number of taxa known to exist at that level was determined. Initially I looked for datasets of three types: (1) those containing all known OTUs; (2) those not containing all OTUs but where the author gave explicit justification for their omission (e.g., uncertain assignment to the group or poor preservation); (3) those not adhering to a specific taxonomic scheme, but sampling OTUs with the express intention of capturing variation in morphology within the clade (Brusatte *et al.* 2012). Datasets with uneven sampling were included provided the OTUs in the dataset could be amalgamated in such a way as to render them homogeneous at some higher taxonomic level. For example, if a dataset contained twenty genera, ten of which were unique exemplars of ten different families and ten of which were exemplars of the same (eleventh) family, then those latter 10 genera were amalgamated to produce a dataset for (effectively) eleven families.

Character amalgamation utilised modal states. If there were equal numbers of character states then one character state was chosen at random. Stratigraphic data were amalgamated using the oldest and youngest exemplars (total range). If the hypothetical eleven families represented all of those known within an inclusive clade (or order, in taxonomic terms), then the dataset was deemed to be effectively complete at the family level. Datasets of between 50% and 100% completeness were included provided that the coded OTUs offered reasonably even coverage of the more inclusive group. Where coverage was less than this or manifestly uneven (unless the author gave good reasons why additional OTUs were not included; e.g., poor preservation, uncertain affinity to the

clade in question), the OTUs were amalgamated to represent the next highest taxonomic level.

Where a dataset contained two clades, each sampled homogeneously but at two different taxonomic levels, it was split into two subclades. For example, the data set of Carlson & Fitzgerald (2007) contained two clades; one of 49 homogeneously sampled genera, the other of 20 homogeneously sampled genera. Where homogeneously sampled clades were nested within larger inhomogeneously sampled clades, OTUs were pruned from the dataset until only the homogenous clade remained. Pruning was required in 26 of the 98 datasets. All pre-analytical modifications to the data were implemented in *Excel*, and all morphological condensing was conducted in the software environment *R* (2012) using a script (Appendix.1) written by Dr Sylvain Gerber (University of Cambridge) and Martin Hughes (University of Bath).

A potential concern was that any putative pattern of disparity profiles through time might result from the analysis of clades of markedly different taxonomic rank. For example, many metazoan phyla originate in the Cambrian and most of those that do not are without a fossil record (Valentine *et al.* 1994). It follows that clades of higher rank and longer duration are more likely to originate in older intervals, whereas clades of genera and families (for example) can potentially originate at any time. Some of this potential bias was mitigated by the stipulation that clades must be entirely extinct; many higher taxa not only originated in older strata but also persisted to the Recent, and were consequently disqualified. From the sample of 98 extinct clades, only three spanned from the Palaeozoic to the Cenozoic. The sampling therefore preferentially selected taxa of intermediate and lower rank (orders and families). The relationship between clade age of origination and clade rank was tested.

Taxonomic ranks are only comparable in the very broadest terms across different metazoan groups. To test for the relationship between taxonomic rank and clades age, I coded subfamilies as rank one, families as rank two, superfamilies as three, up through ranks of infraorders, orders, superorders, subclasses, classes and subphyla at rank nine.

### **1.7 Stratigraphic Resolution**

After effective taxonomic ranks were ascribed to all OTUs in a matrix, stratigraphic ranges were assigned to reflect that ranking. I used the *International Stratigraphic Chart 2009* (Ogg *et al.* 2008; Gradstein *et al.* 2004) offering a maximum resolution of 100 stages. First and last occurrences were determined at the maximum possible precision for each OTU. Taxa were assumed to persist uninterrupted between these dates. Stratigraphic range data were sourced from the Paleobiology Database (Alroy *et al.* 2001), Sepkoski *Online* (Sepkoski 2002), the *Fossil Record 2* (Benton 1993), and from other published sources (Fortey & Chatterton 1988; Novas *et al.* 2009; Pollitt *et al.* 2005; Wang *et al.* 2004; Wang *et al.* 1999; Smith & Kroh 2011; Mitchell 1987; Harper *et al.* 2000; Body



2000; Wilson & Reeder 2005; Fortelius 2011; Eades *et al.* 2011; Ponder & Lindberg 2008; Sprinkle & Kier 1987; Smith 2011; Uetz 2011; Brands 2011; Benton 2000; Carroll 1988; Kroh & Mooi 2011; Evenhuis 1994; Mah 2009; Forrest 2011; Schmidt 2011; Kullman 2011; Bervoets 2011; Brake 2011; Slieker 2011), including the papers containing the morphological data matrices (Appendix.1.Table.2).

Ranges for higher taxa were determined by searching for those taxa directly, *and* by searching for all constituent taxa within them. In general, I coded the maximum duration for a given OTU. If no stratigraphic data could be found for a given OTU (and there was no evidence to suggest that it was extant) then it was noted and removed prior to analysis. Unranked OTUs (*incertae sedis*) were dated individually. Stratigraphic ranges for each OTU were formatted as binary matrices of absences and presences. Data taken from the Paleobiology Database and Sepkoski *Online* were formatted into a stratigraphic matrix using a *Perl* script (written by Anne O'Connor, University of Bath). The script produced the binary matrix by cross-referencing the OTUs in a dataset with a reference list of taxa and dates extracted from both databases. First and last occurrences were then checked manually against other published sources. Clades were plotted according to their first occurrences.

The analysis of all 98 clades was run with temporal bins of three different sizes (stages, epochs and periods). For each of these bin sizes, disparity profiles were categorised. In a small number of cases, it became necessary to utilise more inclusive categories than geological periods. Specifically, the Palaeozoic was partitioned into Early Palaeozoic (Cambrian/Ordovician) and Late Palaeozoic (Silurian – Permian) (Gould *et al.* 1987). Other treatments of the Palaeozoic (retaining a single interval or splitting it into two bins comprising either equal numbers of epochs or with equal durations) were also implemented, but made negligible difference to the overall results. The choice of stratigraphic resolution was otherwise important because not all clades had the same duration (the mean duration of clades (Appendix.1.Table.2) had a mean duration of fifteen stages or 90 million years and a standard deviation of 12 stages or 69 million years).

There was a need to balance between sampling at a higher temporal resolution (i.e. Stage, Epoch, Period) on one hand and sampling more OTUs within each time bin on the other hand. Attempting to use too many stratigraphic subdivisions potentially results in problematic levels of noise, prohibitively large error bars and larger numbers of intervals containing just one or no OTUs while large subdivisions would result in few intervals in which to plot disparity through time. In practice, the preferred binning regime yielded between 5 and 15 intervals. The paucity of top-heavy clades in the Cenozoic was partly a function of using entirely extinct groups. Many datasets of Cenozoic fossil taxa contained extant representatives, and could not be pruned to yield an exclusively extinct subclade.

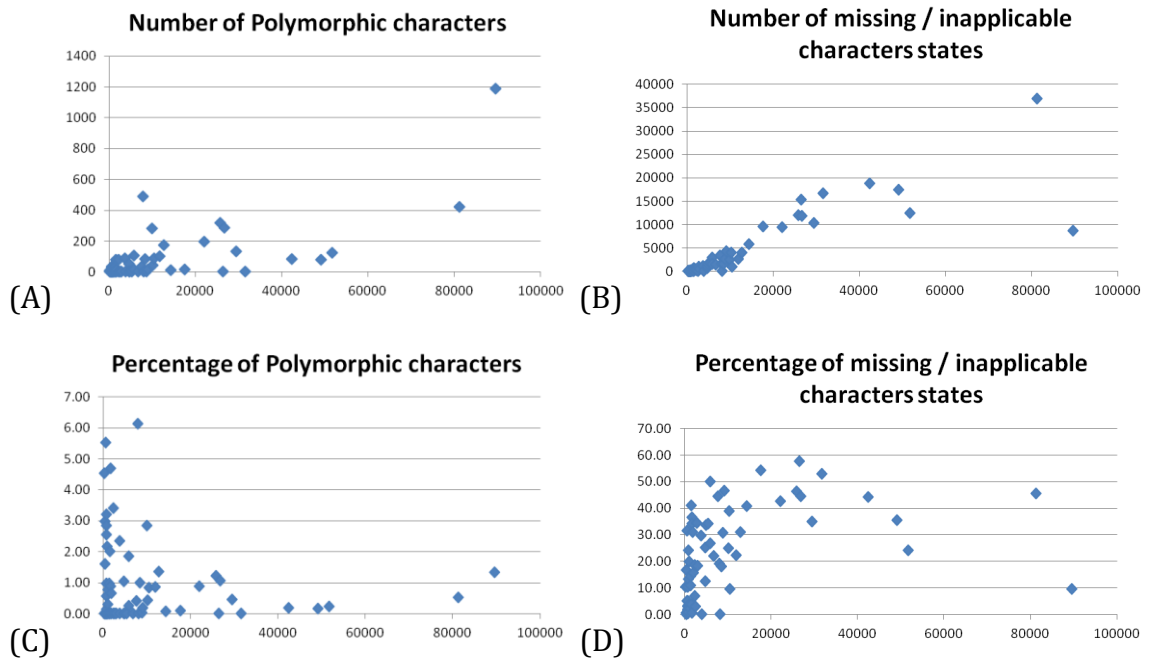


Figure.1.5. Number of problematic characters vs. total number of characters per dataset.. (A) Number of polymorphic character states in the whole morphology matrix per dataset. (B) Number of missing/inapplicable character states in whole morphology matrix per dataset. (C) Percentage of polymorphic characters states as part of the whole morphology matrix per dataset. (D) Percentage of missing/innaplicable characters states as part of the whole morphology matrix per dataset. (Appendix.1.Table.3)

### 1.8 Character Coding

Characters were coded as found in the source paper. If OTUs were to be condensed as a result of uneven sampling, character states were merged in the following way. (1) Where there was no character state conflict resulting from the condensing, no action was taken. (2) If the character states conflicted between OTUs, I assigned the state most frequently observed within the rest of the matrix. A total of 33 studies (36 clades) contained information on the ordering of character data. For these studies the analysis was run to see what effect ordering or not ordering characters had on the disparity curves and centre of gravity estimates.

Polymorphic character codes constituted only 1.5% on average and no more than 7%. As a conservative approach, they were treated as NA prior to condensing taxa (Fig.1.AB). Missing and/or gap state data were more prevalent (Fig.1.5CD), up to a maximum of 70% but on average made up only 20% of datasets. Even at its highest concentration, it was possible to calculate pairwise distances between almost all of the taxa in all matrices. In cases where the condensation of taxa rendered some characters invariant, these were deleted prior to analysis.

### 1.9 Intertaxon Distance Matrices & Ordination

The above procedures resulted in one morphological character matrix and one stratigraphic presence/absence file for each of the 98 clades. All subsequent analyses were conducted within *R* using scripts written by Sylvain Gerber and Martin Hughes. I first

generated a Generalised Euclidean Distance (GED) matrix (Wills 1998b) between all pairs of OTUs. For each pair of OTUs, the difference between each comparable character state was squared, these differences summed, divided by the total number of comparable states, then square rooted (Wills 1998b). A small number of morphological matrices contained one or two pairs of OTUs for which no characters were coded in common (hence it was impossible to calculate a distance between them). The authors of those studies had included them to show that they had considered the range of available specimens. The most data deficient OTUs were removed from the matrix and the distances recalculated (repeated as necessary) until distance measures were produced for all remaining pairs of OTUs in the matrix. Distance matrices were ordinated using principal coordinates analysis (Wills *et al.* 1994) and implementing Cailliez's correction for negative and uninformative Eigen values (Cailliez 1983). Non-metric multidimensional scaling (NMDS) (Anderson *et al.* 2011; Shen *et al.* 2008) was *not* chosen as that method does not retain the metric properties of the distance matrix after ordination (Huntley *et al.* 2006).

### **1.10 Quantifying Disparity**

The sum of variances measure of disparity was used throughout the analyses of this chapter (Foote 1991a). It was chosen because (1) it is one of the most commonly used metrics discrete character disparity analyses (the other being sum of ranges), and (2) it is a simple to calculate and effective measure of the spread of specimens through the morphospace across time. This measure was favoured because of its relative insensitivity to sample size differences and its complete indifference to the orientation of the coordinate axes upon which it is being calculated. The first property is useful because the number of OTUs sampled can be highly variable through time, and because larger samples will tend to have a wider distribution within their character space. Uncorrected measurements of range are sensitive to these changes in sample size, to the extent that they may largely reflect proxies of diversity. The second property is useful because although the orientation of the ordination is derived from all taxa simultaneously, taxa within a given time bin are almost invariably some subset of these. Were these subsets to have been ordinated independently, the orientation of axes would very likely have been different.

The sum of ranges disparity metric (Foote 1991a; Wills 2001; Ruta 2009) was also calculated with every disparity analysis and time period variation, but did not reveal any trends. The reason behind this phenomenon was the level of rarefaction used to adjust the disparity in each time interval to account for the differences in sample size. Rarefaction is commonly used in diversity analyses to take into account differences in taxonomic diversity as some of the variation in absolute diversity through time is likely not a real biological pattern but instead because some organism either were never fossilised

or have yet to be found as fossils. By uniformly subsampling the lowest possible diversity of a clades time series (e.g. Stage 1 has 5 taxa while Stages 2-5 have 10 each) one assumes that the average diversity through time more accurately tracks the changes in diversity than the absolute values of diversity. The same process is used on the sums of ranges disparity metric as it suffers from a similar problem with unsampled diversity. The sum of ranges metric is sensitive to outliers, i.e., genera with morphologies away from the global centroid or the clade morphospace may have not been sampled due to chance. If they had been fossilised the sum of ranges value for any particular interval may have a very different value.

To combat the problem, each dataset was rarefied so disparity in each interval was calculated using the number of taxa from the time interval with the lowest standing diversity. Rarefaction is very harsh on clades with time intervals with few taxa. In the analysis the majority of datasets required rarefaction to 2-3 taxa to include the whole time series. This effectively flattened the disparity value across the time series. One future modification of the analysis would be to choose a higher rarefaction threshold, but what threshold to pick would become arbitrary. A more effective refinement would be to use a different subsampling method such as shareholder quorum subsampling considered by some to be a more suitable subsampling method for fossil diversity calculation (Alroy *et al.* 2008; Alroy 2010a). Disparity was calculated for all time intervals over which each clade persisted. 1000 Bootstrap replicates (Efron 1982) were used at each interval to produce estimates of mean and standard error. Stratigraphic resolution was to stage, epoch or period level depending upon the size and duration of the clade (with between 5 and 15 time intervals per clade).

### 1.11 Analyses

Empirical morphospaces were derived as multidimensional spaces in which the proximity of OTUs correlated with their morphological similarity (Wills *et al.* 1994; Briggs *et al.* 1992). Disparity was measured using the sum of variances on successive axes of the morphospace (Wills *et al.* 1994; Foote 1992a; Foote 1992b). In order to derive a trajectory of disparity through time, the duration of the clades was split into time bins; defined so as to balance the competing requirements of stratigraphic resolution and sample size (Foote 1992b) (Fig.1.4A). To provide a single index of the shape of clade disparity profiles, the centre of gravity (CG) metric previously applied to paleontological diversity and disparity data was calculated (Gould *et al.* 1987; Uhen 1996; Foote 1991a).

The centre of gravity in absolute time ( $CG_m$ ) was given by:

$$CG_m = \frac{\sum_{i=1}^n d_i t_i}{\sum_{i=1}^n d_i} \quad (1)$$

Where  $d_i$  is the disparity at the  $i$ th interval of  $n$  stratigraphic intervals and  $t_i$  is the temporal midpoint in absolute time (millions of years ago) of the  $i$ th interval. I then scaled this value between the ages of the oldest ( $t_{oldest}$ ) and youngest ( $t_{youngest}$ ) representatives of the clade to a scaled index of observed CG ( $CG_{scaled}$ ) between 0 and 1.

$$CG_{scaled} = \frac{t_{oldest} - CG_m}{t_{oldest} - t_{youngest}} \quad (2)$$

If time bins were all of the same duration, then clades with uniform or symmetrical disparity profiles would have  $CG_{scaled}$  of 0.50 (mid-way) (Fig.1.6). Clades with a relatively early disparity maximum ('bottom-heavy') would have  $CG_{scaled} < 0.50$ , while those with a late disparity maximum (top-heavy) would have  $CG_{scaled} > 0.5$  (Fig.1.6).

In practice the expected  $CG_{scaled}$  for a clade of constant disparity through time is not necessarily 0.50, but rather determined by the durations of the time bins over which the profile was measured. This is because stratigraphic stages are of variable durations, and because taxa are not always dated to series and stages. Hence, the  $CG_{scaled}$  was compared with the inherent  $CG_{scaled}$  ( $CG_i$ ) for a hypothetical clade of uniform disparity spanning the same intervals. A bootstrapping test determined when this deviation was significant (clades for which >97.5% of 1000 bootstrapped replicates lay either above or below the centre of gravity inherent in the timescale ( $p$ -value <0.05) (Foote 1991a). Finally, the observed scaled  $CG_{scaled}$  was adjusted relative to  $CG_i$  as a zero baseline; hereafter known simply as CG. Clades were then partitioned into one of three categories according to CG: significantly bottom-heavy, significantly top-heavy and indistinguishable from symmetrical (Fig.1.6). Log likelihood ratio goodness-of-fit tests (G-tests) were used to compare frequencies of different profile shapes (e.g., in different time bins).

Clades that were not significantly top- or bottom-heavy could nonetheless have a variety of profile shapes. To split up these shapes, an ancillary test was used to determine whether the taxa observed at the beginning and end of the history of each clade (those in the first and last time bins) had a disparity pattern that could be distinguished from the maximum observed in any time bin (Fig.1.7). The disparity profile of the clade was resampled using 1000 bootstraps of all the OTUs in the dataset. For each replicate curve, the difference in disparity between the first (or last) intervals and the disparity maximum elsewhere in the curve was calculated, yielding a distribution. If a difference of zero was within the 95% limits of this distribution, I could not reject the null hypothesis of no

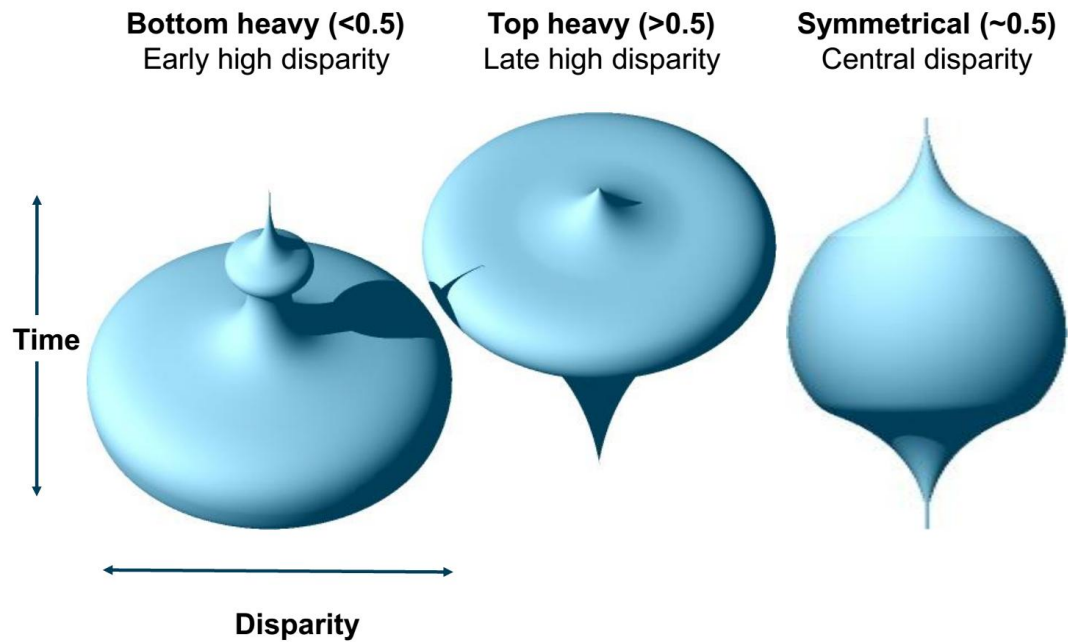


Figure.1.6. Theoretical disparity profiles (filled out in 3D) showing the three centre of gravity categories (Left) Bottom-heavy disparity profiles recognised by CG values significantly below 0.5. (Middle) Top-heavy disparity profiles recognised by CG values significantly greater than 0.5. (Right) Symmetrical disparity profiles recognised by CG values not significantly different from 0.5.

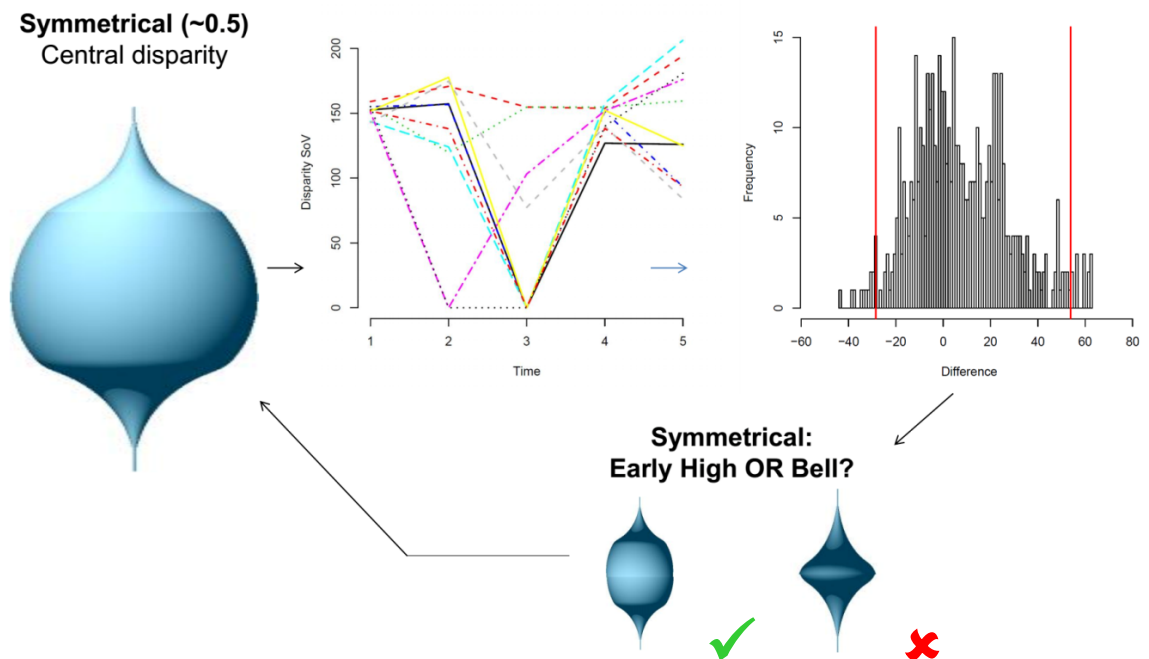


Figure.1.7 Summary of the process of categorising disparity profiles into early high or bell shaped curves. Theoretical symmetrical disparity profile given as an example but process was applied to all disparity curves. Process is as described in the text. The disparity profile of the clade is bootstrapped and differences between the maximum disparity and the first interval disparity are calculated, shown by a frequency histogram. In the above example the process picks early high (green tick) over bell shape (red cross) as the bootstrap distribution contains values of 0 (i.e. no difference between the maximum disparity and the disparity in the first stage).

difference between the initial disparity and the maximum achieved by the clade (Fig.1.7). Combining the ancillary test with the CG test resulted in a total of four categories of disparity profile: significantly bottom-heavy, significantly top-heavy, early saturated and other (Fig.1.7).

For profiles of diversity through time, standard birth/death models of clade evolution on a regular timescale predict a mean CG of 0.5 (Uhen 1996; Gould *et al.* 1977; Kitchell & MacLeod 1988). There is no analogous model for profiles of *disparity* through time. Many clade histories are characterised by early periods of (variously rapid) radiation and morphological diversification, but new forms *must* be generated by branching cladogenesis (and so clades must explore the design space progressively at some level of temporal resolution). The extinction of OTUs *may* also be highly selective, affecting branches in a progressive manner that reflects the structure of the tree. Extinction can also be effectively random with respect to both the tree and the morphospace, which would tend to leave variance based indices of disparity largely unaffected (Foote 1991b).

Any model in which the random extinction of lineages is a significant factor is likely to yield top-heavy clades as a null expectation. The use of an effective CG of 0.5 as the null is defensible in this context because this top-heavy 'bias' would be in favour against finding more significantly bottom-heavy than top-heavy clades. It is also defensible empirically based on the sample. The median CG of the 98 clades (corrected for intrinsic bias) was 0.497, which is indistinguishable from 0.5 ( $V = 2304$ ,  $p = 0.668$ ). Clades can *only* become more diverse through the progressive branching of lineages, but may become less diverse because of extinctions that may be either selective (large branches) or effectively random. It is therefore noteworthy to find clades that exhibit early levels of disparity that are indistinguishable from the maximum disparity later achieved by the clade.

### **1.12 Removing the Immediate Effects of Mass Extinction**

Clades terminating at mass extinction boundaries might be shaped significantly by those extinction events. Specifically, a mass extinction might truncate a clade that would otherwise have obtained a symmetrical or bottom-heavy profile of disparity through time. Such extinction-terminated clades might be expected to be flat-topped to some degree, and therefore have a potentially higher CG than otherwise. To remove the most immediate of these effects, I reduced the sample to include only those clades not terminating at one of the big five mass extinction boundaries; namely the End Ordovician (443.7Myrs), the Late Devonian (Frasnian/Famennian boundary 374.5Myrs), the End Permian (251Myrs), the End Triassic (199.6Myrs), and the End Cretaceous (65.5Myrs). Broader definitions are possible for the End Ordovician and Late Devonian events, but these had no effect upon any of the statistics. A clade was deemed to terminate at a mass

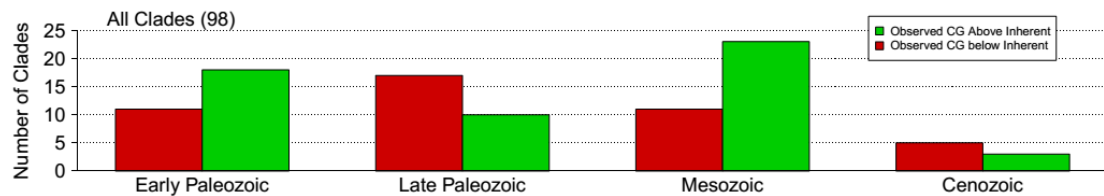


Figure.1.8. Counts of disparity by Era for all 98 extinct clades. The Palaeozoic was split into the Early (Cambrian/Ordovician) and Late (Silurian-Permian). Red are datasets with a CG value <0.5; green are datasets with a CG value >0.5.

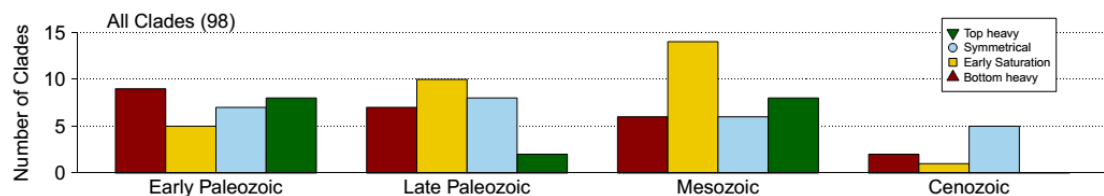


Figure.1.9. Counts of disparity by Era for all 98 extinct clades. The Palaeozoic was split into the Early (Cambrian/Ordovician) and Late (Silurian-Permian). Dark red are datasets with a CG values that significantly lower than 0.5 after 1000 bootstraps; dark green are datasets with a CG value significantly above 0.5. Yellow are those symmetrical datasets that also qualified at early saturation (i.e. maximum disparity was no different from disparity in initial first/second stage). Light blue are those symmetrical dataset that did not qualify as early saturation.

extinction if *at least* one of its constituent OTUs went extinct at the boundary, but none persisted thereafter.

### 1.13 Comparison of 'Cladistic' & 'Phenetic' Sourced Matrices

The majority of the discrete matrices analysed in this study had been used in other contexts to infer phylogeny (hereafter 'cladistic' matrices). A small minority were not prepared with this purpose in mind (hereafter 'phenetic' matrices, in the absence of a better term). Do cladists code characters differently from those *not* intending to derive a tree? In particular, do cladists 'filter' potential characters more stringently than pheneticists, applying more rigorous criteria of operational homology and thereby rejecting sources of variation that would be legitimate within analyses of morphological disparity?

All morphological data sets – whether intended for phylogenetic purposes or otherwise – necessarily sample a subset of the available universe of possible characters. Moreover, many pheneticists utilised concepts of homology for discrete characters that differed little from the *operational* definitions used by cladists (i.e., those of close compositional and relational similarity). Hence, even if rich seams of discrete 'phenetic' data matrices were available in the literature (which, unfortunately, they are not); I believe that these would be largely similar in quality and nature to most cladistic matrices. Most pheneticists did not, for example, code states believed to be blatant analogies (wings of birds, bats and bees)



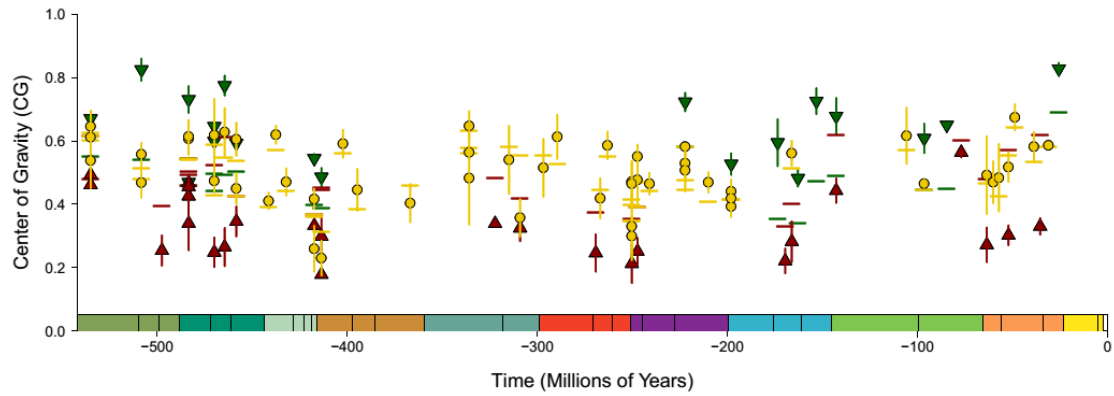


Figure.1.10. Centre of gravity ( $CG_{scaled}$ ) values for all 98 datasets across the Phanerozoic. Case studies are sampled relatively evenly throughout this time, and there is no systemic temporal trend in disparity profile shape. Circles denote mean scaled CG ( $CG_{scaled}$ ) from 1000 bootstrap replicates of the variance-based disparity curves for each clade, plotted against the clade origination date. Vertical lines denote the standard error around  $CG_{scaled}$ , derived from 1000 bootstrap replicates. Green triangles; significantly top-heavy profiles ( $CG_{scaled} > CG_i$  with  $p < 0.05$ ). Red triangles; significantly bottom-heavy profiles ( $CG_{scaled} < CG_i$  with  $p < 0.05$ ). Yellow circles; profile indistinguishable from symmetrical. Abscissa colour scheme: international stratigraphic chart.

as the same state. They might, perhaps, have taken greater licence with characters than *some* cladists, but this is a difference of degree rather than one of principal. By the same token, nearly all cladistic matrices contain homoplasy, and many of them retain characters that map onto globally optimal trees without supporting any non-terminal branches (i.e., they map with the maximum possible number of changes). Such characters are typically retained, even though their homology is implicitly rejected by the rest of the data.

### 1.14 Results

For *diversity* through time, random birth/death models with constant parameters predict that the average clade shape should be symmetrical (Gould *et al.* 1987; Kitchell & MacLeod 1989). For *disparity*, the predictions are less precise. New species can only arise from the fission of existing ones (clades initially diversify from a single species and therefore a single point in morphospace), whereas extinctions can be random with respect to this same tree (Ciampaglio *et al.* 2001). Therefore, if a clade follows a homogeneous birth-death model with Brownian character evolution, some top-heaviness would be expected (Foote 1991a). The use of 0.5 as a null is slightly simplistic but biased against the principal finding (namely that clades not terminating at a mass extinction event are bottom-heavy on average).

Across the 98 clades (including those terminating coincident with a mass extinction) disparity profiles provided a mean CG of 0.495, with a median CG (0.501)

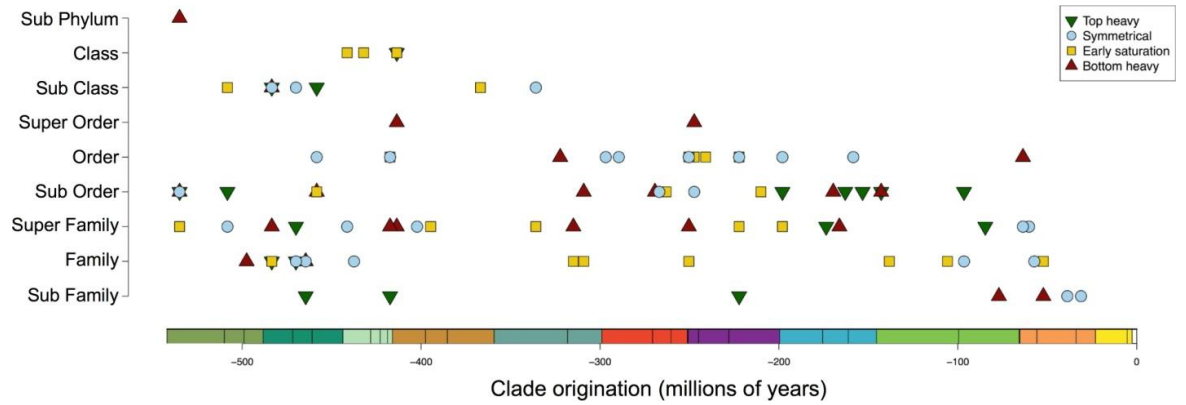


Figure.1.11. Scatter plot of clade rank against time of clade origination. Dark Red; significantly bottom-heavy ( $CG_{scaled} < CG_i$  at  $p < 0.05$ ). Dark green; significantly top-heavy ( $CG_{scaled} > CG_i$  at  $p < 0.05$ ). Yellow: clades not significantly bottom- or top-heavy but display early saturation (ancillary test) in disparity during their evolutionary history. Pale blue: clades not significantly bottom- or top-heavy and do not display early saturation of disparity during their evolutionary history.

indistinguishable from 0.5 ( $V = 2429$ ,  $p = 0.992$ ). Time-averaged indices masked some apparent differences in clade disparity profiles within and between eras; most notably, there were more bottom-heavy ( $CG < 0.5$ ) clades in the Late Palaeozoic than top-heavy ( $CG > 0.5$ ) clades, with the opposite pattern in the Mesozoic (Fig.1.8). Comparison across four time bins (Early Palaeozoic (Cambrian/Ordovician), Late Palaeozoic, Mesozoic and Cenozoic) revealed no significant differences in the frequencies (log likelihood ratio test;  $G = 2.298$ ,  $p = 0.513$ ). After implementing the bootstrapping test for *significant* deviation from clade symmetry, the data was partitioned into three groups: significantly bottom-heavy, significantly top-heavy and indistinguishable from symmetrical (which I discounted) (Fig.1.9). Again, there were no significant differences in the relative frequencies of significantly top- and bottom-heavy clades across the four time bins ( $G = 3.558$ ,  $p = 0.313$ ). Finally, a plot of clade CG against the time of clade origin revealed no systemic trends throughout the Phanerozoic (Fig.1.10).

The next set of results looked at the affect of taxonomic resolution on CG through time. Spearman's rank correlation were computed between taxonomic rank and the age of clade origination; this revealed a weak but significant relationship ( $\rho = -0.213$ ,  $p = 0.035$ ) (Fig.1.11). When the age of the clade (first stage of origination) was changed to the age of the clades CG in absolute time, the correlation between taxonomic rank and age was even weaker and no longer significant ( $\rho = -0.147$ ,  $p = 0.148$ ). The relationship between CG and taxonomic rank was tested but again no significant correlation was found ( $\rho = -0.003$ ,  $p = 0.979$ ) (Fig.1.12A).

When CG was partitioned into familial (sub families, families and superfamilies) ordinal (infraorders, orders and superorders) and higher (subclasses, classes and subphyla) taxonomic groups no difference was found between the median CGs (Kruskal-Wallis test;  $X^2 = 0.603$ ,  $p\text{-value} = 0.740$ ) (Hughes *et al.* 2013). Hence, even if

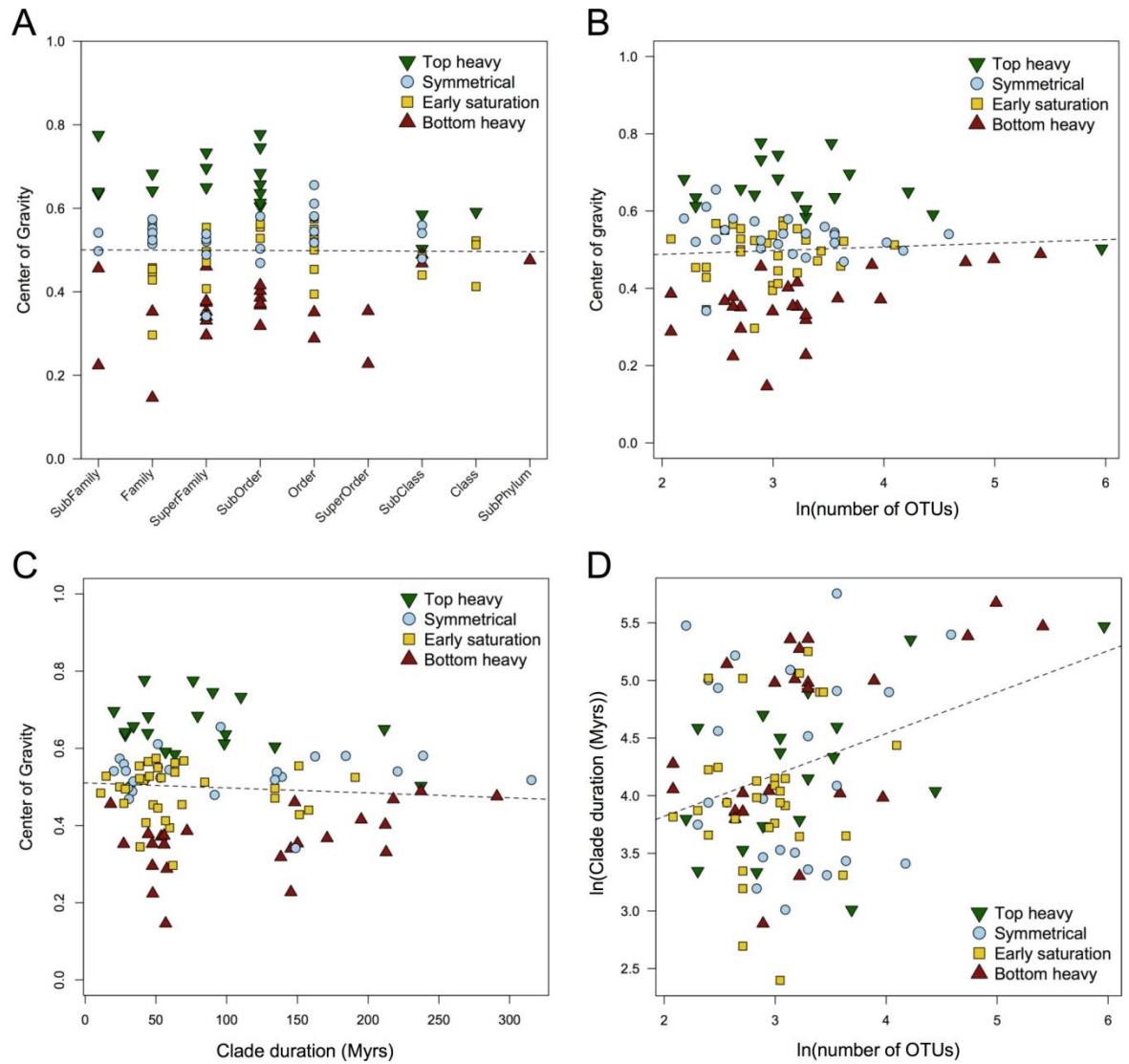


Figure.1.12. There is no correlation between the centre of gravity of clades and either the taxonomic level of their constituent OTUs, the number of constituent OTUs or the duration of the clade. (A) Scatter plot of clade CG against median clade rank ( $\rho = -0.0027$ ,  $p = 0.9791$ ). (B) Scatter plot of clade CG against log(number of OTUs) ( $\rho = 0.0147$ ,  $p = 0.8857$ ). (C) Scatter plot of clade CG against clade duration (Myrs) ( $\rho = -0.0986$ ,  $p = 0.3342$ ). (D) Scatter plot of log(clade duration) against log(number of OTUs) ( $\rho = 0.2241$ ,  $p = 0.0265$ ). Dark Red: significantly bottom-heavy ( $CG_{scaled} < CG_i$  at  $p < 0.05$ ). Dark green: significantly top-heavy ( $CG_{scaled} > CG_i$  at  $p < 0.05$ ). Yellow: clades not significantly bottom- or top-heavy but display early saturation of disparity during their evolutionary history. Pale blue: clades not significantly bottom- or top-heavy and do not display early saturation of disparity during their evolutionary history.

higher-ranking clades were preferentially sampled from older strata, this could not bias disparity profile shapes in and of itself.

There was no significant correlation between clade CG and the number of OTUs as a proxy for clade size (Fig.1.12B), or between clade CG and duration of the clade (Fig.1.12C). There was a significant positive correlation between the logged number of OTUs and logged total duration of the clade ( $\rho = 0.2241$ ,  $p = 0.0265$ ), i.e., older and longer a clade has been around for, the more taxa it accumulates. This was not be a problem as the addition of genera does not necessarily result in an increase in disparity (Wesley-Hunt 2005) Finally, CG was modelled as a linear function of log(number of

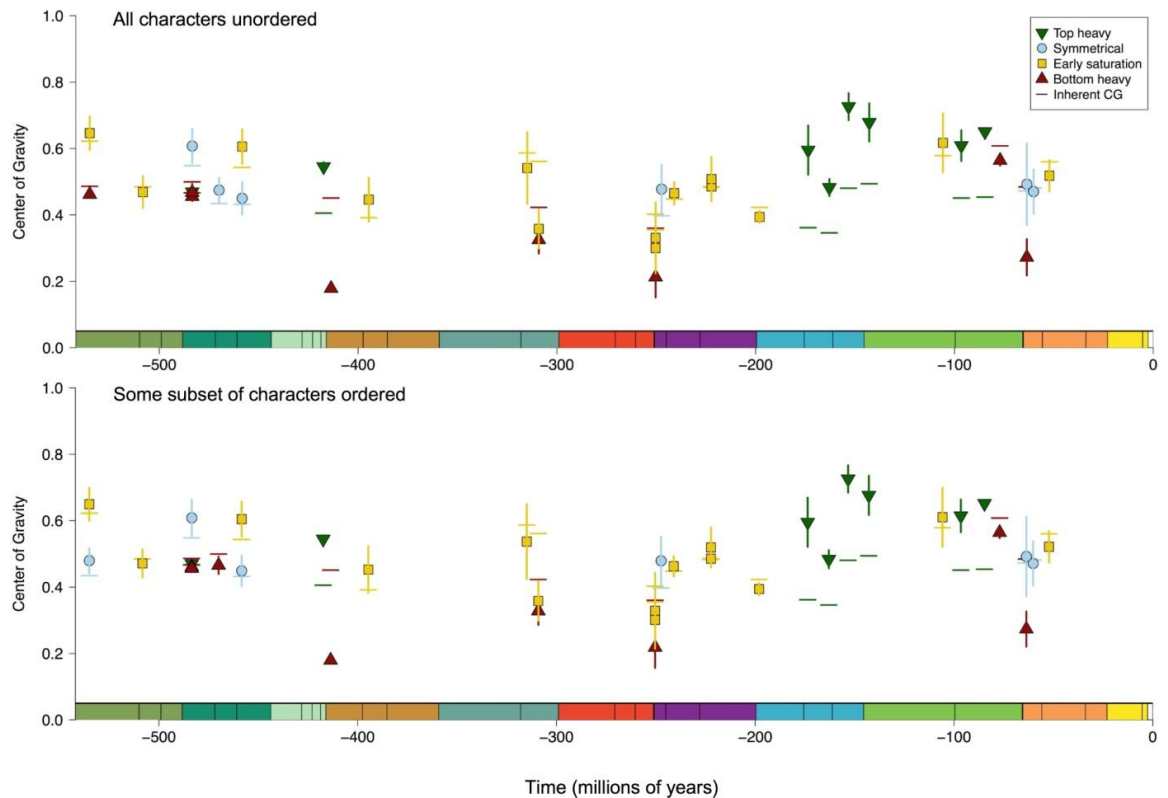


Figure.1.13. Ordering character states has a very small effect upon disparity calculations. Of the 98 clades analysed, 36 included information on the ordering of one or more characters. For this subset of 36 clades, the analysis was run with all character states unordered (upper panel), and with some states ordered as specified in the original publication (lower panel). Dark Red: significantly bottom-heavy ( $CG_{scaled} < CG_i$  at  $p < 0.05$ ). Dark green: significantly top-heavy ( $CG_{scaled} > CG_i$  at  $p < 0.05$ ). Yellow: clades not significantly bottom- or top-heavy but display early saturation of disparity during their evolutionary history. Pale blue: clades not significantly bottom- or top-heavy and do not display early saturation of disparity during their evolutionary history.

OTUs) + log(taxonomic level of clade) + log(duration of clade in stages) + date of clade origin + taxonomic group (vertebrate or invertebrate) + demise of clade at mass extinction (yes or no). The only significant variable (and the only variable retained in the minimum adequate model using the Akaike information criterion (AIC) (Johnson & Omland 2004);  $P < 0.001$ ) was the last of these; whether or not a clade went extinct coincident with a mass extinction (Hughes *et al.* 2013).

The following results looked at the effect that character ordering could have on the potential disparity values (36 datasets in all). Figures 1.13 to 1.15 demonstrate that these differences were undetectable. The Figure 1.13 plots the 36 datasets CG values through time via their origination point. There is no visible trend in either the unordered or ordered data through time. The Figure of 1.14 shows the frequency of each CG category totalled across time. There were no differences in the numbers of CG categories tallied over the Phanerozoic. Figure 1.15 shows the disparity trajectory of athyridid brachiopods (Fig.1.15A) and birds from the Spheniscinae (Fig.1.15B). The disparity trajectories of both were only slightly affected by character ordering, the same results were seen across other disparity datasets, with no obvious pattern in the direction of the shift upward or

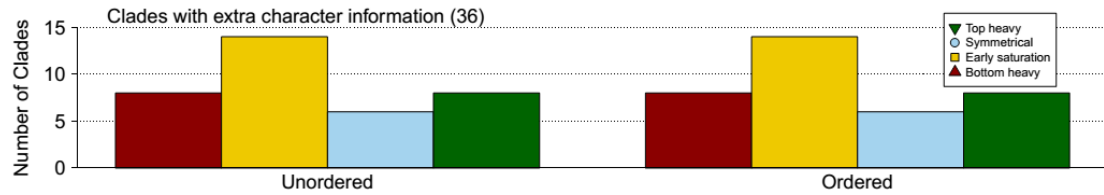


Figure.1.14 Counts of datasets that have used unordered vs. ordered characters states. Dark Red: significantly bottom-heavy ( $CG_{scaled} < CG_i$  at  $p < 0.05$ ). Dark green: significantly top-heavy ( $CG_{scaled} > CG_i$  at  $p < 0.05$ ). Yellow: clades not significantly bottom- or top-heavy but display early saturation of disparity during their evolutionary history. Pale blue: clades not significantly bottom- or top-heavy and do not display early saturation of disparity during their evolutionary history.

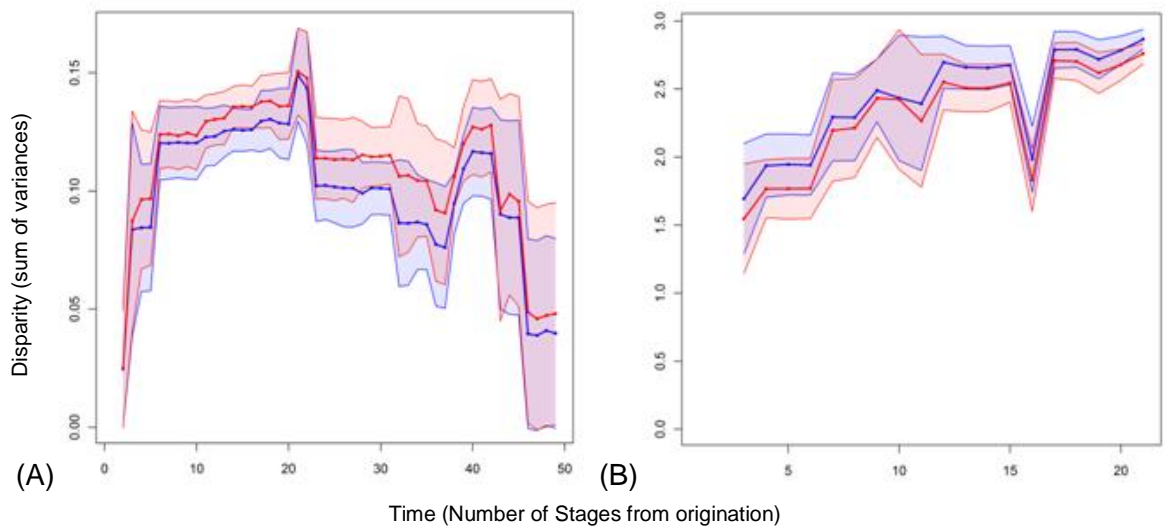
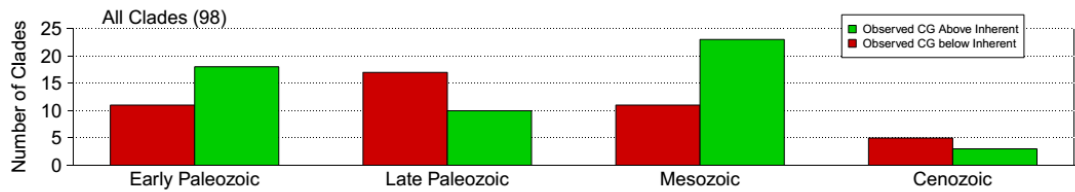


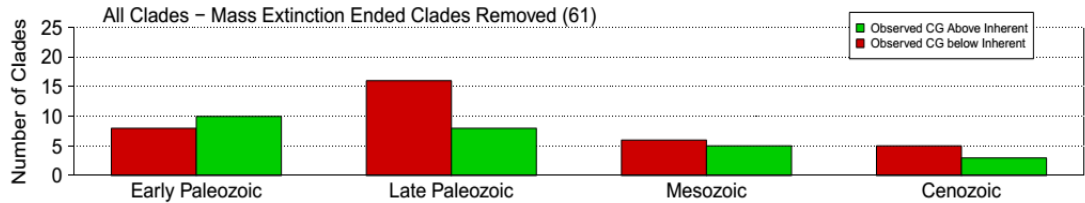
Figure.1.15. Comparison of the effect that ordered vs. unordered character data has on individual disparity profiles. (A) Disparity through time (Dapingian-Tithonian) of athyridid brachiopods (Alvarez *et al.* 1998). (B) Dapingian-Tithonian (Selandian-Holocene) of Spheniscinae, Aves (Bertelli *et al.* 2006). Thick lines are the mean of observed variance at each time bin, thin lines are the upper and lower bounds of standard error for each time bin. Shaded area indicates the level of overlap between ordered and unordered data. Blue: unordered character data; Red: ordered character data.

downward. The correlation of  $CG$  values (corrected for intrinsic bias) for ordered and unordered analyses across the 36 clades was extremely high (high ( $\rho = 0.950$ ,  $p\text{-value} = < 2.2e-16$ )). Unordered data were therefore used for all further analyses, and are the basis of the results presented elsewhere (Chapters 2-4).

Frequencies of top- and bottom-heavy clade profiles in the four largest time bins (Early Palaeozoic, Late Palaeozoic, Mesozoic and Cenozoic) were compared, irrespective of whether these were significant or not (Fig.1.16A). A log likelihood ratio test applied to a 2 x 4 contingency table of all 98 clades retained the null of similar distributions within these bins ( $G = 7.353$ ,  $X^2$   $df = 3$ ,  $p = 0.061$ ). I then counted *only* those clades for which the

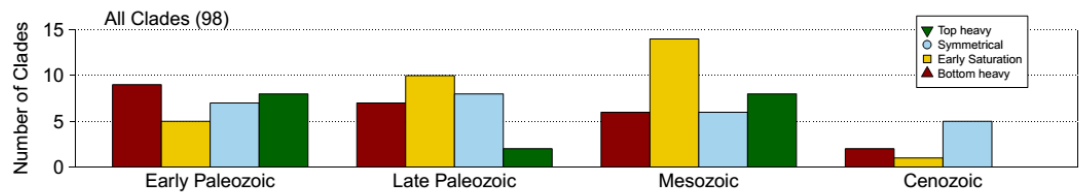


(A)

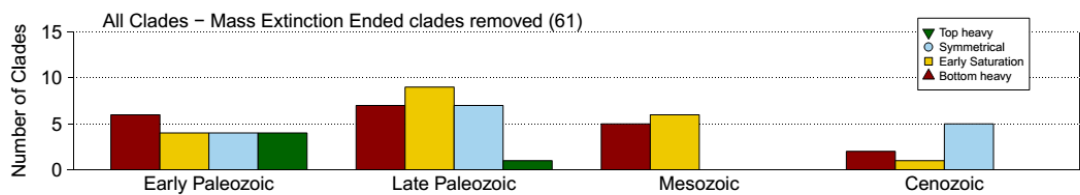


(B)

Figure.1.16. Counts of CG in each Era. Palaeozoic is split into the Early Palaeozoic (Cambrian-Ordovician) and Late Palaeozoic (Silurian-Permian). Red are those CG values  $< 0.5$ ; green are those CG values  $> 0.5$ . (A) All 98 clades; (B) only those clades not ended at mass extinctions.



(A)



(B)

Figure.1.17. Counts of disparity CG category by Era. The Palaeozoic is split into the Early (Cambrian-Ordovician) and Late Palaeozoic (Silurian-Permian). (A) All 98 extinct datasets. (B) All 61 datasets not ending at mass extinctions. Dark Red: significantly bottom-heavy ( $CG_{scaled} < CG_i$  at  $p < 0.05$ ). Dark green: significantly top-heavy ( $CG_{scaled} > CG_i$  at  $p < 0.05$ ). Yellow: clades not significantly bottom- or top-heavy but display early saturation of disparity during their evolutionary history. Pale blue: clades not significantly bottom- or top-heavy and do not display early saturation of disparity during their evolutionary history. Mass extinctions: Late Ordovician, 443.7Myrs; Late Devonian, 374.5Myrs; Late Permian, 251Myrs; Late Triassic, 199.6Myrs; Late Cretaceous, 65.5Myrs.



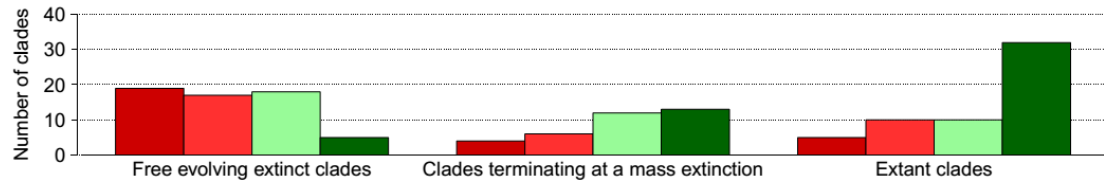


Figure.1.18. Groups terminating at one of the ‘big five’ mass extinction events (and living groups that are still diversifying) are more top-heavy than those terminating at other times. (Left) Disparity profile frequencies for extinct clades that terminate at mass extinction boundaries. (Middle) Disparity profile frequencies for extinct clades that do not terminate at mass extinction boundaries. (Right) Disparity profile frequencies for living clades (truncated by the Recent). Dark Red: significantly bottom-heavy ( $CG < 0.50$  with  $p < 0.05$ ). Dark green: significantly top-heavy ( $CG > 0.50$  with  $p < 0.05$ ). Light red and light green bars indicate the frequencies of clades with  $CG$  below and above 0.50 respectively, but for which  $p \geq 0.05$ . Mass extinctions: Late Ordovician, 443.7Myrs; Late Devonian, 374.5Myrs; Late Permian, 251Myrs; Late Triassic, 199.6Myrs; Late Cretaceous, 65.5Myrs.

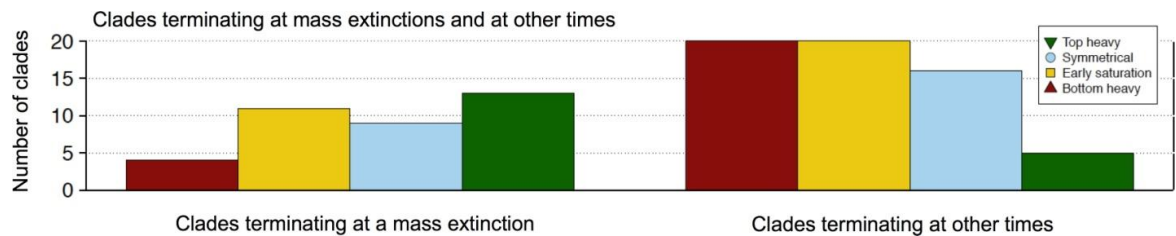


Figure.1.19. Clades terminating at one of the ‘big five’ mass extinction events are more top-heavy than those terminating at other times. (Left) Disparity profile frequencies for clades that terminate at mass extinction boundaries. (Right) Disparity profile frequencies for clades that do not terminate at mass extinction boundaries. Dark Red: significantly bottom-heavy ( $CG_{scaled} < CG_i$  at  $p < 0.05$ ). Dark green: significantly top-heavy ( $CG_{scaled} > CG_i$  at  $p < 0.05$ ). Yellow: clades not significantly bottom- or top-heavy but display early saturation of disparity during their evolutionary history. Pale blue: clades not significantly bottom- or top-heavy and do not display early saturation of disparity during their evolutionary history. Mass extinctions: Late Ordovician, 443.7Myrs; Late Devonian, 374.5Myrs; Late Permian, 251Myrs; Late Triassic, 199.6Myrs; Late Cretaceous, 65.5Myrs.

bootstrapping test allowed us to reject the null hypothesis of symmetry, yielding another 2 x 4 contingency table albeit with a smaller total sample size (1.17A). Again, a log likelihood ratio test retained the null of similar distributions within the four time bins ( $G = 5.2$ ,  $X^2$  df = 3,  $p = 0.158$ ). Finally, when all 98 clades were partitioned into four groups (‘significantly high  $CG$ ’, ‘significantly low  $CG$ ’, ‘early saturation’ and ‘other’), distributions in the four time bins again showed no significant deviation from the null ( $G = 16.750$ ,  $X^2$  df = 9,  $p = 0.053$ ) (Fig.1.17A).

Although clade disparity profiles had a mean  $CG$  indistinguishable from 0.5, there was a marked and significant difference in  $CG$  between those clades terminating coincident with a mass extinction event and those becoming extinct at other times (Fig.1.18). Of the

37 clades that terminated at a mass extinction, 24 had a  $CG > 0.5$ , whilst the median  $CG$  of the 37 clades was significantly greater than 0.5 (one sample exact Wilcoxon test;  $V = 537$ ,  $p\text{-value} = 0.004$ ) (i.e., were 'top-heavy' on average) (Fig.1.16B). Of the 61 clades that terminated at other times (Fig.1.16B), just 21 had a  $CG > 0.5$ , and the mean of all 61  $CG$ 's was significantly less than 0.5 ( $t = -2.657$ ,  $p = 0.010$ ), making non mass extinction terminated clades bottom-heavy on average. A likelihood ratio test revealed that these profile frequencies were significantly different (2 x 2 contingency table, log likelihood test;  $G = 8.127$ ,  $X^2 df = 1$ ,  $p = 0.004$ ) (Hughes *et al.* 2013). A similar difference was observed for significantly top- and bottom-heavy clades (Fig.1.17B). Twelve from 17 significantly asymmetrical clades terminating at a mass extinction were top-heavy, while just 5 from 25 significantly asymmetrical clades terminating at other times were top-heavy (Fig.1.19) ( $G = 11.074$ ,  $X^2 df = 1$ ,  $p = 0.000$ ) (Fig.1.19).

Unsurprisingly, the two distributions also had significantly different medians (Wilcoxon rank sum test;  $W = 1637$ ,  $p = 0.000$ ). The clades becoming extinct at other times than mass extinction events had significantly 'bottom-heavy' disparity profiles on average (63 clades with a mean  $CG$  significantly less than 0.500;  $t = -2.420$ ,  $p = 0.018$ ). By contrast, the 35 clades ending at mass extinctions had a mean  $CG$  significantly *greater* than 0.500 ( $t = 3.901$ ,  $p < 0.001$ ) (Fig.1.19). Likelihood ratio tests also confirmed that the relative frequencies of top- and bottom-heavy clades terminating at mass extinctions and at other times were different, whether including all clades ( $G = 7.648$ ,  $p = 0.006$ ) or only those with significant skew ( $G = 13.022$ ,  $p < 0.001$ ) (Hughes *et al.* 2013).

I also tested to see if for those clades that ended at mass extinctions, and those that were bounded by mass extinctions whether there was a pattern of  $CG$  changing depending on distance from the mass extinction. Figure 1.20 shows the results of those taxa that ended at a mass extinction (Fig.1.20A) and those that were bounded by mass extinction (Fig.1.20B). If the  $CG$  value of disparity was reliant on the clades distance from mass extinctions (i.e. top-heavy clades found close to mass extinction events, bottom-heavy clades directly after them, symmetrical clades in between) then a pattern should be discernible. In both cases there was no pattern of a transition on disparity patterns as distance from mass extinction events decreases. The data here required pooling all mass extinctions together, therefore some patterns may yet be discernible for specific extinction events if sample sizes were increased and so the mass extinctions could be looked at individually.



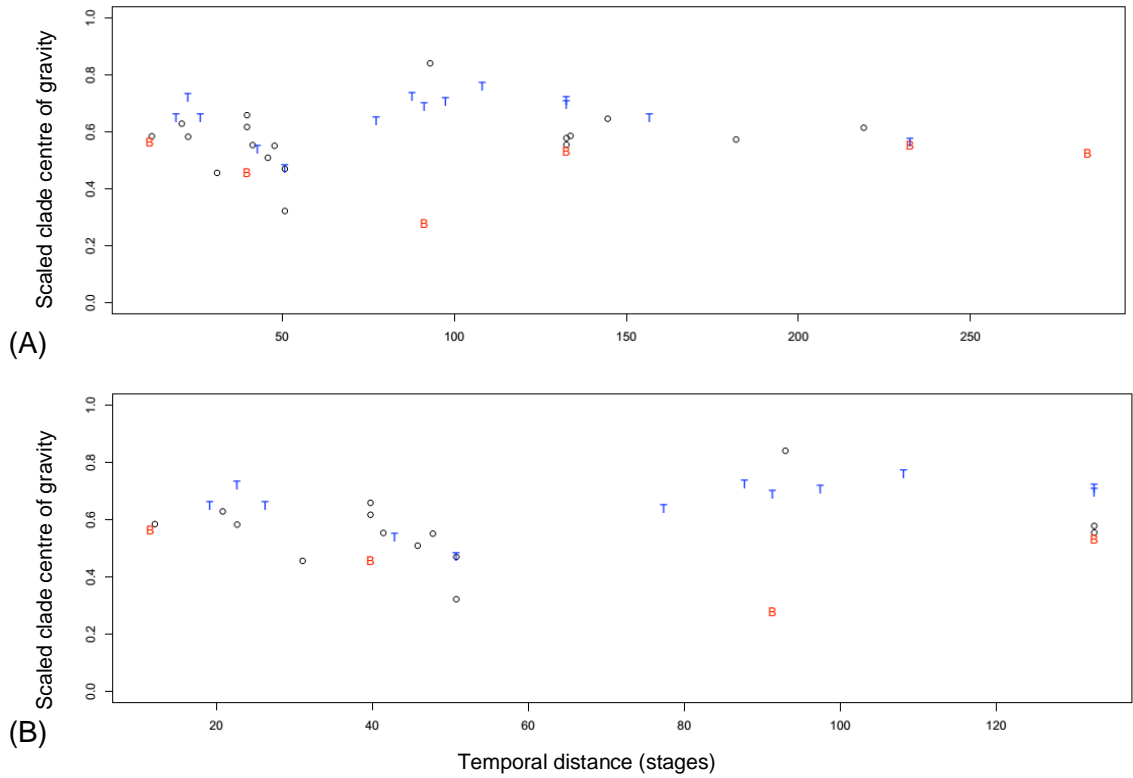


Figure.1.20. Temporal distance (number of stages) of the scaled clade CG from mass extinction event, i.e.0 distance is the actual extinction event, 10 stages equals 10 stages from mass extinction event. All mass extinction events are pooled to increase sample size. (A) All mass extinction ended clades; (B) all mass extinction ended clades that are bounded by mass extinctions events (begin in the wake of one and become extinct at another). Red B: significantly bottom-heavy  $CG_{scaled}$ ; blue T: significantly top-heavy  $CG_{scaled}$ ; black circle: non-significant  $CG_{scaled}$  values (symmetrical).

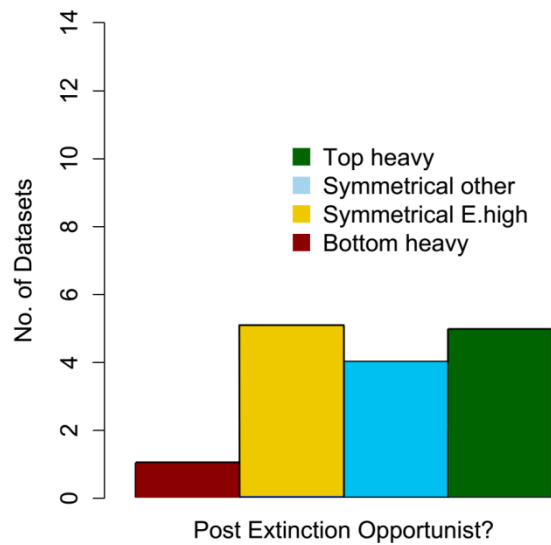


Figure.1.21. The disparity CG categories for those clades evolving in the wake of mass extinctions. Dark Red: significantly bottom-heavy ( $CG_{scaled} < CG_i$  at  $p < 0.05$ ). Dark green: significantly top-heavy ( $CG_{scaled} > CG_i$  at  $p < 0.05$ ). Yellow: clades not significantly bottom- or top-heavy but display early saturation of disparity during their evolutionary history. Pale blue: clades not significantly bottom- or top-heavy and do not display early saturation of disparity during their evolutionary history.

I also looked at the disparity of those clades that ended at mass extinctions. These datasets did not show an early burst in disparity (significantly early high disparity) as expected, an instead found an equal number of the other disparity categories (Fig.1.21). A similar suite of tests was run on only that subset of 61 clades that did not terminate at a mass extinction event. None of these revealed significant differences between Early Palaeozoic, Late Palaeozoic, Mesozoic or Cenozoic time bins, irrespective of how the data were partitioned. The partition into clades with CG above and below 0.5 ( $G = 2.098$ ,  $X^2$  df = 3,  $p = 0.552$ ) and with CGs significantly above and below 0.5 ( $G = 4.397$ ,  $X^2$  df = 3,  $p$ -value = 0.222) revealed no significant differences. The partition into four classes ( $G = 17.649$ ,  $X^2$  df = 9,  $p$ -value = 0.034) revealed one or more significant differences.

Taken together, these results imply that differences in the distribution of clade profile shapes in the four time bins are a function of differential impact of the mass extinctions at the ends of these bins upon the sample. Most notable is the high frequency of significantly top-heavy clades terminating at the end of the Mesozoic. Indeed, these are the only top-heavy clades in the era. For comparison, I also generated disparity profiles for 53 *additional* living clades with high diversity in the Recent (Appendix.1.Table.2) (these clades are otherwise excluded from the final conclusions about disparity trends through time unless stated). These extant clades (truncated by the present) had a median CG significantly greater than 0.500 ( $V = 1150$ ,  $p < 0.001$ ), but which was indistinguishable from that for fossil clades terminating at a mass extinction ( $W = 924$ ,  $p = 0.980$ ) (Fig.1.18). Over half of the study clades had disparity profiles that were neither significantly top- nor bottom-heavy. These 'symmetrical' clades may nonetheless have a variety of trajectories, with their own particular macroevolutionary implications. Most remarkable are groups (e.g., crinoids (Foote 1994a)) whose earliest exemplars have levels of disparity that are not significantly different from the maximum levels subsequently achieved by the clade; a simplistic null of early maximal disparity. For 29 of the 54 'symmetrical' groups, I were unable to reject this null. Such a pattern would be close to that often envisaged for explosive radiations (Gould 1989; Gavrilets & Losos 2009) and similar to that proposed as the trajectory for metazoans through the Phanerozoic (Wills & Fortey 2000).

Early high patterns inevitably imply an unsampled period of cladogenesis (or the existence of ghost ranges) at the base of the clade, but this either occurs too fast for the available stratigraphic resolution or is not fossilised (Wills & Fortey 2000; Fortey *et al.* 1996). *Late* saturation is much less remarkable, since clades have already undergone radiation and diversification, and had almost the entirety of their histories in which to colonise the extremities of their morphospaces. While late saturation was observed in 32 'symmetrical' clades, twelve of these also ended at a mass extinction (and were therefore likely to have been prematurely truncated).

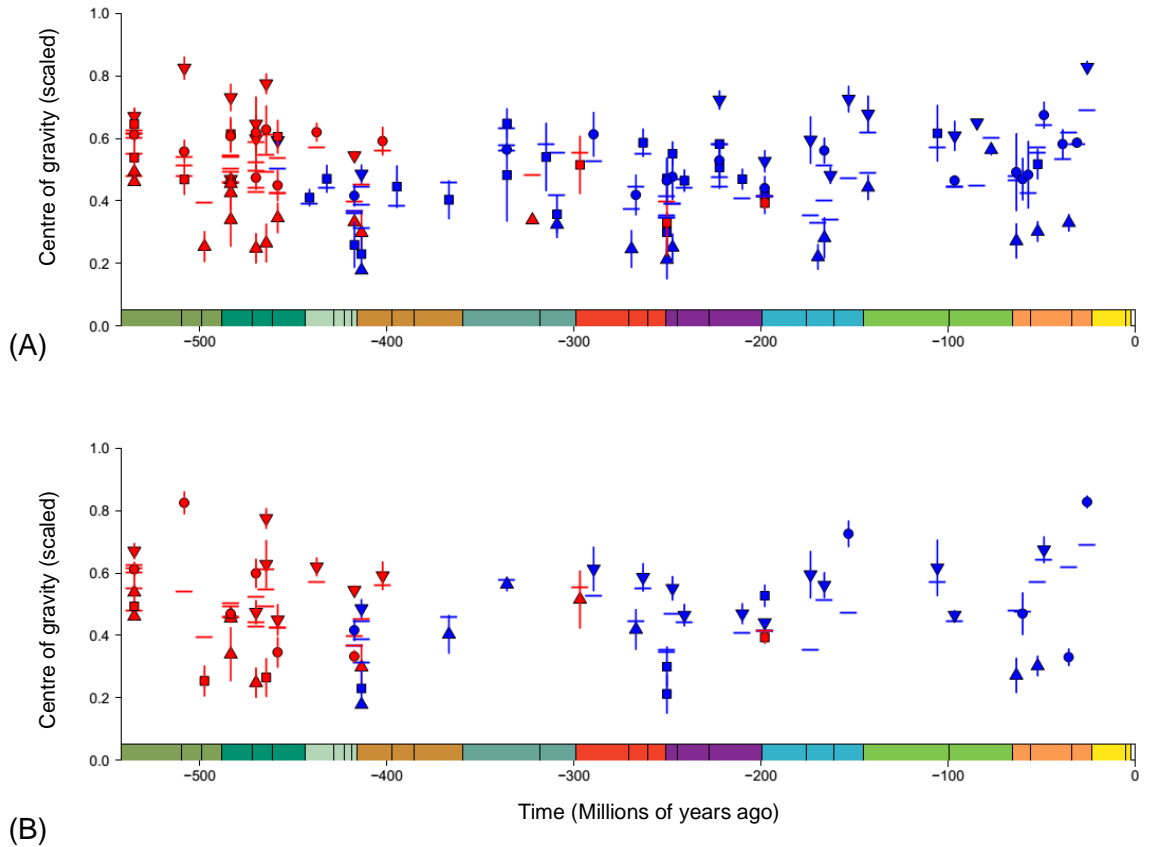


Figure.1.22. Distribution of vertebrate vs. Invertebrate groups ( $CG_{scaled}$ ) across time across. (A) All 98 extinct clades. (B) All 53 extant clades. Red indicate invertebrate clades. Blue indicate vertebrate clades. Symbols indicate the type of CG profile of the clade. Downward pointing triangle: top-heavy,  $CG > 0.5$  at P-value  $< 0.05$ . Upward pointing triangle: bottom-heavy,  $CG < 0.5$  at P-value  $< 0.05$ . Square: early saturation,  $CG < 0.5$  at P-value  $> 0.05$ . Circle: symmetrical,  $CG \approx 0.5$  at P-value  $> 0.05$ .

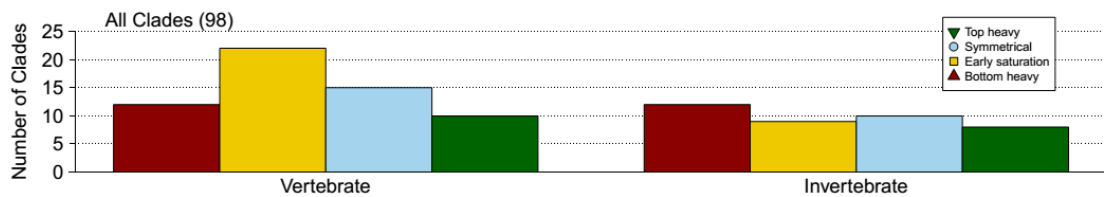


Figure.1.23. Counts of CG profiles of all 98 extinct clades split by (Left) invertebrate, and (Right) vertebrate grouping. Dark Red; significantly bottom-heavy ( $CG_{scaled} < CG_i$  at  $p < 0.05$ ). Dark green; significantly top-heavy ( $CG_{scaled} > CG_i$  at  $p < 0.05$ ). Yellow; clades that are not significantly bottom- or top-heavy but display early saturation of disparity during their evolutionary history. Pale blue; clades that are not significantly bottom- or top-heavy and do not display early saturation of disparity during their evolutionary history.

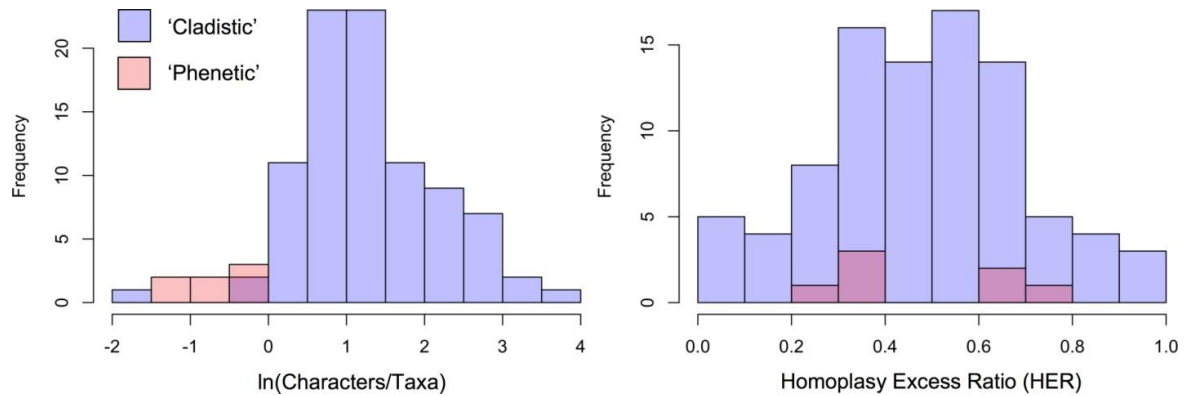


Figure.1.24. The distributions of data matrix 'shape' (the ratio of the number of characters to the number of taxa) and homoplasy (indexed here by the homoplasy excess ratio; HER) for the 91 'cladistic' (blue) and 7 'phenetic' (pink) matrices used in this study. 'Cladistic' matrices are defined as those compiled with the express purpose of inferring a phylogeny, whereas 'phenetic' matrices are those compiled without this purpose in mind.

For this reason, I again focused on the 63 'free evolving' clades that did not terminate at a mass extinction. Of these, the proportion (two thirds) that were either significantly bottom-heavy or showed early saturation (two mutually compatible conceptions of 'early high disparity') was significantly greater than the proportion that were either significantly top-heavy or showed late saturation ('late high disparity') (two-sample test for equality of proportions;  $\chi^2 = 4.613$ ,  $p = 0.016$ ). These results suggest clades that do not terminate at a mass extinction do indeed *tend* to reach their highest levels of disparity relatively early in their evolutionary histories (Gould 1991). Moreover, the results find that this tendency occurs throughout the Phanerozoic rather than being a pattern most common in the Palaeozoic (Gould 1989).

I split the data into those clades there were invertebrate or vertebrate across time and tallied up across the Phanerozoic to see how disparity was distributed across these two major groupings. There is a rough temporal split (with some overlap in the middle Phanerozoic) in the data between the vertebrate and invertebrate clades through time (Fig.1.22). This does not affect the overall conclusions of the results herein which find early high disparity as a prevalent trend across the Phanerozoic as seen in Figure 1.23. There are also more extinct vertebrate clades in my sample but this does not represent a major bias in the data. Many invertebrate clades that have evolved have not yet become extinct and so were included in the extant partition of the data (Fig.1.22B).

The temporal split of invertebrate datasets into mostly the Palaeozoic partly reflects the fact that many of the known invertebrate phyla evolved early in the evolutionary history of the Earth (Valentine *et al.* 1999) and so are expected to have originations in the Palaeozoic half of Figure1.22. The temporal split may also partly reflect an artefact of scientific effort and interest. There is the possibility that there are simply more available discrete character datasets of Palaeozoic invertebrates than Mesozoic or Cenozoic. The analyses herein made every effort to sampling fairly across time, but even with 98 extinct

and 53 extant clades the sample is still a very modest and too small to test whether the latter possibility is true, but could be a potential avenue for future research.

The 'cladistic' or 'phenetic' nature of the character matrices were also looked at to see if this would bias the disparity results (Fig.1.24). The former datasets are often typified by a higher ratio of characters to taxa than phenetic matrices, which is borne out in the very modest sample here (Fig.1.24). One could go a step further, and include functional characters. Such data have greater utility for investigating functional variety rather than morphological variety and disparity as usually conceived. Some datasets were excluded from the analyses for this reason.

To compare the two kinds of datasets, the small number of 'phenetic' datasets were analysed as though they were intended for phylogenetic inference, subsequently calculating the homoplasy excess ratio (HER) (Archie 1989a; Archie 1989b; Archie 1996) for these and the other matrices in the study. The HER is an index designed to find if the observed character states along any tree are comparable to those generated from random data that are phylogenetically comparable, i.e., data with the same characters states per character but randomised across taxa. Values close to 1 suggest no homoplasy while values close to 0 suggest all characters have no phylogenetic information. The HER is relatively unbiased by dataset dimensions unlike the ensemble consistency and retention indices that test for homoplasy by assessing the minimum and maximum number of character changes along the tree (Hammer & Harper 2006).

Calculating the HER requires an estimate of the mean tree length for a large number of datasets (here 500) in which the assignment of states within characters (but across taxa) has been randomised. For this, Dr Matthew Wills (University of Bath) used a *TNT* (137) script implementing *xperm* followed by *xmult level 10* searches. The latter (in the absence of further *bb* searches) is very unlikely to find the full set of optimal trees, but is guaranteed to find some of them (and hence the true optimal tree length, which is all that is needed within the HER calculation). The results indicate no difference in the median HER for 'cladistic' and 'phenetic' data sets ( $W = 323$ ;  $p = 0.917$ ). Hence, the 'cladistic' matrices sampled greater numbers of characters (relative to the number of taxa in the matrix) than their 'phenetic' counterparts. Coded 'cladistic' and 'phenetic' characters are indistinguishable in terms of their conflict with one another and the amounts of homoplasy that they introduce. Hence, there is no evidence in the results presented here, given the admittedly modest sample, that cladists use more stringent criteria of operational homology, causing them to reject characters that others might accept (Hughes *et al.* 2013).

One last analysis was conducted to see if the morphological characters themselves had reached some exhaustion point (Gavrillets 1999) which could potentially be used as an explanation for early high disparity. To this end, a simple procedure was used to find whether the morphological distances between genera in each dataset were

the same as the theoretical maximum distances that could be achieved given the character data for those datasets. Where all characters are unordered, the maximum theoretical Euclidean distance between taxa was given simply by the square root of the number of characters. This was compared with the distribution of  $(N^2-N)/2$  empirically observed distances between  $N$  taxa to determine whether any observed intertaxon distances reached the theoretical maximum.

This is not a full test of character exhaustion, for which there are better methods available (Wagner 2000; Ruta *et al.* 2006). Instead the exploration here was used to find whether disparity had reached its theoretical maximum between those genera in each time bin rather than the theoretical maximum for the entire clade. This is because taxa could migrate through the morphospace in a way that at no individual time bin would contain taxa as distant as the theoretical maximum. Of the 98 “extinct” and 53 “living” matrices, only 2 (Adrain *et al.* 2008; Hoffmann 2010) contained an empirical distance as great as the theoretical maximum, a rough measure. In neither of these cases were the maximum empirical distances observed between taxa within the same time bin. One of these had a significantly top-heavy CG, and the other was symmetrical. In these two cases, at least, character saturation (if it occurs) did not yield a low CG. I conclude that character state saturation does not seem to be a significant problem but this should be followed up by more rigorous analyses in future (Wagner 2000; Ruta *et al.* 2006). All results are found in Appendix.1.Table.2.

### **1.15 Discussion**

The analyses in this section found that early high disparity denoted by top-heavy and early saturation datasets were the most common patterns across the Metazoa for both vertebrate and invertebrate groups, but the asymmetry of early high disparity in the Palaeozoic compared to the rest Phanerozoic as proposed by Gould (1989) was not found. What might explain the prevailing pattern of early high disparity in clade evolution (Erwin 1994; Wagner 2010)? Both ecological and developmental explanations have been proposed, and the results remain consistent with both. The ‘empty ecospace’ model predicts that clades will radiate and diversify more rapidly when colonising a new environment. This colonisation may occur because ecospace has been vacated by other occupants (e.g., in the wake of some other extinction; typically the result of external, physical factors), or because a hitherto inaccessible environment or other resource has been rendered viable by the acquisition of some novel, ‘key’ adaptation (Simpson 1944; Etienne & Haegeman 2012; Dumont *et al.* 2012) or series of characters (Donoghue 2005) (an intrinsic, biological trigger).

One may expect the wake after mass extinction events to be a perfect place to find early high disparity caused by the emptying of ecospace; however, those datasets that included clades which originated in the first stage after mass extinctions did not show a

preference for any particular type of disparity pattern (Fig.1.21). As for the idea of some key adaptations, it is noteworthy that major clades are often distinguished from their progenitors because they possess distinct and defining sets of derived characters, or because an extant crown is defined relative to an extinct stem. These divisions into a clade and its residual ancestors would otherwise often be arbitrary. For example, rather than delimiting a clade of Aves from within the non-avian theropod dinosaurs, it would be *possible* to define a clade of Aves *plus* some arbitrary 'depth' of theropod dinosaurs. Birds are defined in the manner they are because they acquired a distinctive suite of apomorphies pertaining to the evolution of theropod derived flight; key innovations, in this case, that also enabled them to exploit a new environment. These shifts in anatomy, physiology, behaviour and ecology may themselves explain the differential survival of crowns and stems. The results of early high disparity were based on those clades that were fully extinct.

More generally it is likely that global shifts in climate, sea level and ocean chemistry (coupled with the elevated rates of extinction and turnover that these phenomena engender (Hannisdal & Peters 2011; Mayhew *et al.* 2012; Benton 2009)) have affected the availability of ecospace throughout the Phanerozoic. The only temporal pattern in disparity profile shapes detected in the data was the significant tendency towards top-heaviness in those clades terminating coincident with a mass extinction (which predominantly result from physical drivers (Peters 2008)). Mass extinctions need not increase the subsequent availability of ecospace, but may actually cause its collapse (Erwin 2001). The absence of any systemic trends in clade disparity patterns through time, or of any increased propensity for early high disparity in clades radiating in the immediate wake of mass extinctions (Fig.1.21) suggest that *if* ecological mechanisms have a role, then this is more likely to be mediated via key innovations (which can evolve at any time) and the opening up of new adaptive zones rather than from ecospace clearing.

The other possible explanation could be the reduction in 'evolvability', i.e., the reduction in genetic diversity with specialisation and time (Valentine 1995); however, molecular studies have found that ancestor genomes, with origins spanning back to the Cambrian or even Precambrian, likely contained many of the genes found in modern taxa, despite those ancestor taxa having body plans different from modern body plans (Marshall & Valentine 2010). The loss of genetic 'evolvability' is also hindered by the evidence that even taxa that are highly specialised ecologically have the potential to evolve back towards more generalist roles (Prendini *et al.* 2010). There is also the ability of clades themselves to shape the environment around them, the so called ecosystem engineer role, that needs to be considered in addressing early high disparity (Sousa *et al.* 2009). This dynamic of clades themselves altering the environment to suit their needs, and thus affecting the environments of other clades may play an important role in the clades ability

to evolve new forms. If these clades can successfully modify the environment, there then becomes less reason to establish new morphotypes to occupy new environments, and so after initial colonisation and early burst of disparity, disparity rates may drop.

It must be stressed that ecological and developmental explanations for early high disparity are not mutually exclusive (Erwin 1994). The results here do not allow one to distinguish between these hypotheses. The hypothesis of increasing developmental constraint predicts that the increasing complexity and interdependence of ontogenetic processes with evolutionary time effectively lock down the potential for subsequent morphological innovation (Gould 1989; Arthur 2004; Wallace & Farrow 1999; Smith *et al.* 1985; Kauffman 1985; Wimsatt 1986). Such mechanisms purportedly explain why body plans become invariant and inflexible with time, although mechanisms by which these constraints may be lifted have been posited (Wagner 2011). Notable examples are the tetrapod pentadactyl limb (early tetrapods explored a range of higher digit numbers (Clack 2006)), the seven cervical vertebrae of all mammals except sloths and manatees (otherwise invariant from mice to giraffes (Galis 1999)) and the diagnostic head segmentation of arthropod subphyla (Cambrian genera explored numerous alternatives with relative freedom (Gould 1989; Wills *et al.* 1995)). Such body patterning characters are usually controlled by *Hox* genes, which are also frequently exapted for other (often functionally and positionally unrelated) developmental roles (Chipman 2001). This increasing pleiotropy (more and more varied roles for the same regulatory genes) may account for the observed reduction of developmental lability. Testing this hypothesis would require detailed ontogenetic data far beyond the scope of this study.

### **1.16 Conclusions & Future Prospects**

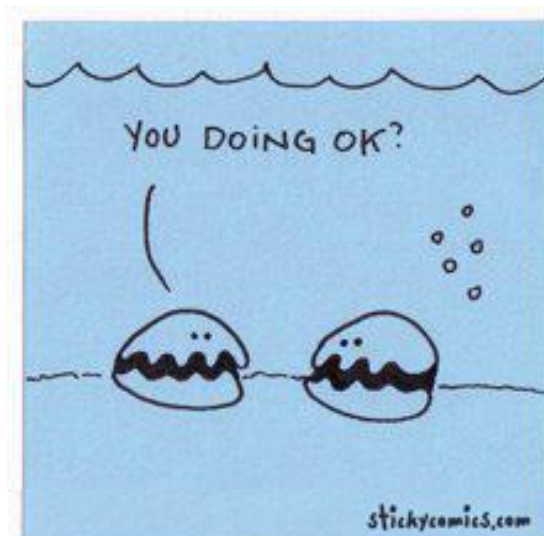
The prevalence of early high disparity as the dominant pattern of clade evolution ranks alongside well-known tendencies for increasing complexity and diversity (Alroy *et al.* 2008; McShea 1996; McShea 1998; McShea & Brandon 2010; Valentine *et al.* 1994) underpinning putative macroevolutionary trends of the widest possible generality. Moreover, it appears to apply throughout the Phanerozoic, and not merely at times of global diversification (e.g., the early Palaeozoic). The distance from mass extinctions and general taxonomic groupings (vertebrate/invertebrate) do not provide any other trends in the data; however, mass extinctions do seem to have some effect as many clades ended at them have top-heavy profiles.

The data here was able to explore the role of characters exhaustion on the data in a limited way, looking at theoretical maximum distances based on characters in individual clades. Character exhaustion is a promising direction of future research using the data of this chapter. Developmental data would also be useful for understanding the degree to which developmental constraints affect clades ability to expand in new directions of morphospace and increase in disparity (Gerber *et al.* 2008).



## ***Chapter 2***

### ***The Diversity of Bivalves Across the Phanerozoic***



## 2.1 Introduction

While Chapter 1 focused on the evolution of disparity patterns across the Metazoa through the Phanerozoic to test if “early high” disparity can be considered an evolutionary “rule” (Hughes *et al.* 2013), this chapter is the first of three that will focus on a specific clade to address variation in levels of diversity and disparity through time and space. I selected the Bivalvia (one of the extant clades in the previous analysis) (Hughes *et al.* 2013) for several reasons. The data is attractive for analysis because it is one of the largest datasets analysed (253 individual genera), and the largest morphological matrix for bivalves ever assembled (Carter *et al.* 2000; Carter *et al.* 2006). The data were provided by Dr Joseph Carter (University of North Carolina). As part of collaboration between Martin Hughes (University of Bath), Dr Matthew Wills (University of Bath) and Dr. Carter some of the analyses of this thesis will contribute to a revised edition of the Treatise of Invertebrate Palaeontology that is currently *in prep.*

Bivalves have a great track record as a clade useful for studying macroevolution (Krug *et al.* 2009; Jablonski *et al.* 2013). Bivalves have much to offer to the study of macroevolutionary history. They are a highly diverse modern clade with an evolutionary history and excellent fossil record that stretches back to the beginning of the Ordovician. They are survivors of every mass extinction to date and have explored an array of aquatic habitats and ecological roles (Ponder & Lindberg 2008). They form the second largest clade of molluscs. Despite their mostly sessile nature, they are found globally, and able to profoundly influence their environment (Sousa *et al.* 2009). They have a body form unique among the molluscs: a soft body encased within two calcified valves. Within these valves they use siphons to draw in water that is then filtered by gills. These gills act as both as their means of absorbing oxygen from the water for respiration but also in many species are used to actively filter food particles from the suspension that are passed to the gut for digestion (Ponder & Lindberg 2008). Their morphological evolution is contradictory as they have conserved the same basic body form since their origination in the Cambrian but also evolved a wide range of morphological structures within this conserved framework.

Many authors have used bivalves, in combination with other taxa or in isolation, as a model group for quantifying changes in Phanerozoic diversity (Miller & Sepkoski 1988; Sepkoski 2002; Jablonski *et al.* 2003; Ros & De Renzi 2005; Kidwell 2005; Dmitriev & Nevesskaja 2006; Alroy *et al.* 2008; Nevesskaya 2008; Vinarskia *et al.* 2011), testing for possible biological and/or taphonomic drivers for these changes (Kosnik *et al.* 2011; Aberhan *et al.* 2012), investigating origination and extinction dynamics (Foote 2000; Bambach *et al.* 2004; Roy, Hunt & Jablonski 2009), identifying trends and testing models of evolution across latitude (Valentine *et al.* 2013; Krug *et al.* 2009; Valentine & Jablonski 2010; Jablonski *et al.* 2006; Roy *et al.* 2000; Roy, Hunt, Jablonski, *et al.* 2009; Roy & Goldberg 2007; Vilhena *et al.* 2013), and assessing the implications of climate change (Watson *et al.* 2012).

Yet basic questions about bivalve evolution remain: is the current diversity trend of bivalves across the Phanerozoic an artefact of old data and old methods or a robust biologically meaningful pattern? How did bivalve disparity change through time? What contribution do constituent subclades make to bivalve diversity and disparity? Does disparity change with latitude like diversity? The aim of the collaboration is clear to use the Paleobiology Database (Alroy *et al.* 2001; Kiessling *et al.* 2013) in combination with a revised edition of the Treatise of Invertebrate Palaeontology (*in prep*) and discrete morphological character data (Carter *et al.* 2000; Carter *et al.* 2006) to test the robustness of current estimates of bivalve diversity across the Phanerozoic (current Chapter), explore how diversity and disparity in bivalves has changed across the Palaeozoic (Chapter 3), and if bivalve disparity changes with latitude (Chapter 4).

## **2.2 Diversity of Bivalves Across the Phanerozoic**

Bivalvia is a Class within the Phylum Mollusca (Ponder & Lindberg 2006). There are over 9000 confirmed extant species (Huber 2010) within >1800 genera across >100 families (Cox *et al.* 1971). This is a working estimate; another >11 000 potential species names lie buried in the palaeontological literature (Huber 2010). They form an ancient clade, originating in the Early Cambrian (542-528 million years ago) (Elicki & Gursu 2009). Compared to their present abundance, size variability, ecological and taxonomic success; Cambrian bivalves were few, small, difficult to distinguish from each other and limited in ecological scope. Only five bivalve genera are recognised during this period: *Arhouriella*, *Camya*, *Fordilla*, *Pojetaia*, *Tuarangia* (Elicki & Gursu 2009). The Ordovician period is where they begin to rise in taxonomic (Cope 2004) and ecological diversity along Gondwanan marine basins (Sanchez 2008). Bivalve families were a minor part of the total marine fauna compared to another superficially similar clade of animals: the Brachiopoda; which were more abundant and taxonomically diverse across the Palaeozoic (Fraiser & Bottjer 2007). Despite their smaller abundance, bivalves could dominate in some Palaeozoic communities and settings (Kriz 2004; Miller 1989)

Bivalves did seem to go through extinction events at the end of the Ordovician and Devonian periods (Cope 2004; Miller & Sepkoski 1988; Vinarskia *et al.* 2011; Raup & Boyajian 1988). The diversity in bivalves from the Early Devonian to the End Palaeozoic was slow, growing exponentially after the End Permian mass extinction up to present day (Ros & De Renzi 2005; Vinarskia *et al.* 2011; Miller & Sepkoski 1988; Dmitriev & Nevesskaja 2006; Jablonski *et al.* 2003). This mass extinction event was a pivotal moment in bivalve evolutionary history. From the End Permian onwards, bivalves became the dominant clade over brachiopods in taxonomic diversity and abundance up to the present day (Fraiser & Bottjer 2007; Gould & Calloway 1980).

Direct competition between bivalves and brachiopods across the Phanerozoic does not seem to have been the cause of this switch. Prior to the event bivalves were

growing in diversity while brachiopods remained stationary (Gould & Calloway 1980). The event itself reset initial diversities in favour of bivalves which then continued to grow while brachiopod diversity remained stationary (Gould & Calloway 1980). Through the Triassic, bivalve diversity rises, exploring epifaunal niches left by brachiopods of the past, and previously unexplored infaunal niches (Ros *et al.* 2011; Stanley 1968). The bivalves are affected again as the end Triassic (Ros *et al.* 2011) and end Cretaceous mass extinctions (Vilhena *et al.* 2013), then grow in diversity up to the present.

The general consensus at the generic level is that bivalves diversified quickly in the Ordovician (Cope 2004; Sanchez 2008), remained stable across most of the Palaeozoic, and rose in diversity up to the present day after the End Permian mass extinction event (Fraiser & Bottjer 2007). Higher level taxonomic patterns show a more linear trend in the growth of bivalve diversity across the Phanerozoic (Ros & De Renzi 2005), but the five big mass extinctions are still visible.

### **2.3 Problems with Current Bivalve Diversity Estimates**

The main aim of the analyses in this chapter is to address the robustness of the Phanerozoic diversity estimates of Bivalvia to alternative methodologies and the effects of sampling bias. In order to address those questions, a more fundamental question needs to be answered first: to what extent do the three major sources of generic stratigraphic ranges i.e. the Sepkoski Database, the Paleobiology Database, and Treatise of Invertebrate Palaeontology, imply congruent patterns? This question is important; the published edition of the Treatise concerning Bivalvia is over 4 decades old (Cox *et al.* 1971). It makes up the bivalve contingent of the Sepkoski Database; data that contributed to the first attempts of a robust, detailed diversity curve for the entire Phanerozoic (Sepkoski 2002; Miller & Sepkoski 1988). Since these published works many new discoveries and revisions to the bivalve fossil record have been made. This old data has been superseded by another database: the Paleobiology Database. This database built upon the data of Sepkoski, and has been the main resource for many published studies of bivalve (and more general marine Phanerozoic diversity) curves (Ros & De Renzi 2005; Vinarskia *et al.* 2011; Miller & Sepkoski 1988; Dmitriev & Nevesskaja 2006; Jablonski *et al.* 2003; Alroy 2010c; Alroy *et al.* 2008).

There are several good reasons for this; the database itself is free, online, and is undergoing constant revision. It does have its share of problems, making the comparison with the three databases necessary. The database is incomplete, and of uneven quality for new and revised entries. The reasons are practical; the database requires a significant contribution of effort, time and manpower to create and revise entries. There are 153 official contributors to this database, but each have academic careers and priorities outside of the remit of this database, and so can only volunteer a limited amount of time to

its upkeep. Where additions or revisions have been made, they will have been applied preferentially according to the interests and expertise of those editing.

These problems do not suggest that the database is not useful, it is by far the greatest open resource of paleontological data available; but these problems must be kept in mind whenever the database is used. It is certainly one of the best data resources of its kind and will inevitably get better with time as more of the scientific community contribute to its upkeep. The revision of the Treatise database by Dr. Joseph Carter (University of North Carolina) offered the ideal opportunity to investigate the patterns implied by a new compilation of stratigraphic data under unified editorial control, and to contrast this with the patterns implicit in the (effectively) crowd sourced Paleobiology Database, and Sepkoski Database that remains unaltered since its publication online (Sepkoski 2002).

Another problem of current diversity estimates, and therefore the understanding of bivalve diversity through time is the method of diversity construction. Although numerous studies have provided updated revisions of bivalve diversity across the Phanerozoic, each of these have used the same method of diversity curve construction i.e. traditional counts (number of taxa per time interval) (Miller & Sepkoski 1988; Jablonski *et al.* 2003; Ros & De Renzi 2005; Kidwell 2005; Dmitriev & Nevesskaja 2006; Nevesskaya 2008; Vinarskia *et al.* 2011). It is now widely acknowledged that the fossil record cannot be taken at face value, requiring some correction for sampling or methodological bias (Alroy *et al.* 2008; Alroy 2010a; Barrett & McGowan 2009; Benton *et al.* 2013; Foote 2000; Lloyd *et al.* 2011; Lloyd *et al.* 2012; Lloyd & Friedman 2013; O'Connor *et al.* 2011; Peters & Foote 2001; Smith & McGowan 2007; Smith & McGowan 2011).

These proxies such as outcrop area, published paper counts, etc, have been the subject of scrutiny for well over half a century (Sheehan 1977; Peters & Foote 2001; Crampton *et al.* 2003; Smith & McGowan 2007; Barrett & McGowan 2009; Benton *et al.* 2013). The way in which the majority of these sampling proxies affect estimates of diversity in bivalves is unknown. Taphonomic or compositional bias can lead to misinterpretations in fossil classification (Lamsdell *et al.* 2013) or phylogeny (Sansom *et al.* 2011). For the moment bivalve diversity estimates seem robust to these types of bias (Kosnik *et al.* 2011; Kidwell 2005), as the majority of characters separating bivalve fossil species pertain to valve morphology (Carter *et al.* 2000; Cox *et al.* 1971). Another possible bias is the so called “pull of the recent”. This is the accumulation of stratigraphic ranges in younger intervals of time resulting from range extensions between modern taxa and their fossil counterparts compared to those extinct taxa with no modern equivalent. Like taphonomic and valve composition bias, the “pull of the recent” effect does not seem to have a great impact on bivalves diversity (Jablonski *et al.* 2003).

Methodological issues are also an important consideration to any diversity curve. Most stratigraphic ranges of fossils naturally overlap, making counting taxa in any time interval problematic as several counting regimes are available: counting every taxon,

counting only those who enter or leave an interval, or counting those exclusive to each interval (Foote 2000). These methods are problematic as they assume that the fossil record either is a reflection of the true diversity through time and/or uniformly biased across time. As not all organisms are fossilised, counting the fossil record directly to estimate diversity could be problematic. Some organisms have better fossil records due to the environment and sedimentary conditions of the habitat they occupied while alive while others had structures that are more easily fossilisable. Finally random chance also plays some role in whether a fossil is found to be added to the diversity estimate or is destroyed/not found and never counted. Resampling or modelling procedures are an alternative to counting methods.

Three resampling methods are available: rarefaction (Miller & Foote 1996), shareholder quorum subsampling (Alroy 2010a), and an unnamed method that assigns stratigraphically wide ranging taxa randomly chosen single interval ranges (Lloyd *et al.* 2011). Each method tries to correct for uneven sampling of the fossil record in a different way. Rarefaction assumes even sampling, i.e., sampling the same number of individuals in each time interval, is the correct way to counter the perceived unreliability of the fossil record. Shareholder quorum subsampling assumes fair sampling, i.e., sampling taxa in each interval until a certain value under the frequency distribution curve has been detected. The third method assumes longer ranges are due to imprecise dating or lumping of taxa, and so randomly assigns single interval ranges to those taxa with longer ranges. This method in particular is new and not tested on a wide range of datasets (Lloyd *et al.* 2011; Lloyd *et al.* 2012). As there is currently no agreed identifier, I shall call it the range uncertainty method.

Another aspect of Phanerozoic bivalve diversity that has not been greatly explored is to what extent do various bivalve subclades contribute to the total bivalve diversity curve across the Phanerozoic? The Bivalvia is a large and diverse class of invertebrates, made up of two major groupings: the subclass Protobranchia, and the subclass Autobranchia. The latter subclass is the larger of the two, making up a significant proportion of the modern bivalve generic diversity (Carter *et al.* 2011; Mikkelsen 2011). Protobranchia, considered the more basal clade (Sharma *et al.* 2013) based on a number of morphological characters (Carter *et al.* 2000; Carter *et al.* 2006). They have hinge teeth that are small and simple in shape (paleotaxodont) and gills used exclusively for respiration (protobranchiate).

The Autobranchia differ from the Protobranchia in using the gills in respiration and feeding (autolamellibranchiate), and in having more complex hinge teeth (heterotaxodont) (Carter *et al.* 2000). In Heteroconchia, the anterior shell margin is discoidal and the muscle scars are of equal size, while Pteriomorpha have stronger posteroventral obliquity, discontinuous fibrous ligament ontogeny and multiple ligament grooves (Carter

*et al.* 2000). Protobranchia are considered the basal taxa (Cope 1999), but the three clades all make their first appearance in the bivalve Ordovician radiation.

The exponential increase of bivalves from the Cambrian to the present is currently considered a real biological pattern. Whether this pattern is real has yet to be adequately addressed using the most recent techniques for diversity calculation. I chose to test the robustness of the generic Phanerozoic bivalve diversity to the effect of three methodologies of diversity curve constructions: traditional counts, shareholder quorum subsampling, and modelling. I also apply the range uncertainty method in selected cases to compare this new technique to the more established resampling methods. As sampling proxies, I apply two commonly used sampling biases: number of formations and collections (Peters & Foote 2001; Frobisch 2008; Butler *et al.* 2009; Mannion *et al.* 2011; Lloyd 2011).

Generic diversity was chosen over higher or lower taxonomic diversity for several reasons. Firstly the generic level provides the largest pool of data to work with i.e. three databases of generic dates. Secondly the curves produced here can be directly comparable to the current literature (Miller & Sepkoski 1988; Jablonski *et al.* 2003; Kidwell 2005; Dmitriev & Nevesskaja 2006; Nevesskaya 2008; Vinarskia *et al.* 2011). Thirdly species level records are less complete than the generic record and only available from one database (Alroy *et al.* 2001; Kiessling *et al.* 2012). Fourthly higher taxonomic rankings such as families are still undergoing revision (Carter *et al.* 2011).

How the major subclades change and contribute to Phanerozoic diversity is also unknown. To amend this knowledge gap, I will also provide working estimates for the change in diversity of these subclades using the information contained within the Paleobiology Database (Alroy *et al.* 2001; Kiessling *et al.* 2013). As the Autobranchia is such a large group I have decided that this clade will be split into its two constituent infraclasses: the Heteroconchia and Pteriomorpha (Carter *et al.* 2000; Carter *et al.* 2011; Bouchet *et al.* 2010) (Carter *et al.* 2000; Carter *et al.* 2011). These three subclades will later be used for diversity/disparity analysis of morphological data provided by Dr Joseph Carter (University of North Carolina) (Carter *et al.* 2000; Carter *et al.* 2006).

## **2.4 Materials & Methods**

### **2.5 Quantifying Diversity**

Diversity was estimated for all bivalves, and each constituent subclade, across the Phanerozoic. Stratigraphic range data were acquired from three separate sources. The first source of stratigraphic information (generic ranges) was supplied by Dr Joseph Carter (University of North Carolina) as part of collaboration between Martin Hughes (University of Bath), Dr Carter and Dr Matthew Wills (University of Bath). The collaboration was agreed upon on the 7<sup>th</sup> of April, 2012. The data itself is also being used by Dr Carter for upcoming chapters in an *in prep* revision of the Treatise of Invertebrate Palaeontology of

Bivalves. The final edition of Treatise dates used in all relevant analyses of this thesis was created on the 27<sup>th</sup> June, 2013.

The second source was the Paleobiology Database (Alroy *et al.* 2001; Kiessling *et al.* 2013). The generic range of Phanerozoic bivalves was compiled several times to adjust for any possible changes made to the data. The first download was made on the 27<sup>th</sup> of the October, 2011. The final download used to construct generic ranges for all Paleobiology Database was made on the 28<sup>th</sup> of May 2013. During this time generic ranges did not change in any noticeable way. Finally, generic ranges were downloaded from the online version of the Sepkoski Database on the 24<sup>th</sup> of September, 2013 (Sepkoski 2002). The unedited generic data available from each source was: Sepkoski 11169 entries, Paleobiology Database 153262 entries, and 4569 ranges from the Treatise.

Diversity was calculated using three different methods. The first method utilised traditional counts of genera. Multiple counting regimes were considered: counting those genera that cross the base of stage boundaries ( $N_b$ ), those that cross the top ( $N_t$ ), those that range through to any part of the bin ( $N_r$ ), those only found in specific bins ( $N_s$ ), summing all observed genera in each stage (SIB), and using the change in genera between successive time bins (first differences) (Foote 2000; Alroy 2010a). To allow comparisons with other case studies of diversity through time (Dmitriev & Nevesskaja 2006; Jablonski *et al.* 2003; Kidwell 2005; Miller & Sepkoski 1988; Nevesskaya 2008; Ros & De Renzi 2005; Vinarskia *et al.* 2011), I decided to use SIB as the traditional counting method as it was the simplest to calculate and the other methods did not offer a substantially different version of events detected using SIB (Foote 2000; Alroy 2010a).

As an alternative to traditional counting methods, subsampling standardisation methods were also implemented as it has been discussed that counting methods can be subject to some level of bias (Alroy 2010a). Two methods were considered: Rarefaction and Shareholder Quorum Subsampling (SQS) (Alroy 2010a; Alroy 2010c). Both require knowledge of the frequency to which taxa occur within individual samples of time. For example: interval A contains 100 taxa, five taxa have an occurrence of 5; another taxon has an occurrence of 75. Interval B has five taxa each with an occurrence of 20. The occurrences (frequency) of A and B would be 100 and 200 respectively.

Both methods subsample to some quota, traditionally the lowest value of occurrences i.e. 100. This is where the methods diverge. Rarefaction subsamples (e.g. 1000 iterations) 100 taxa at random from the pool of possible taxa in each interval and takes the average number of unique taxa returned. Rarefaction assumes uniform sampling is accurate sampling. In interval A, rarefaction would pick up the most common taxon more often than the 5 unique individuals, while in interval B it would pick up each of the five taxa represented at least once.



Shareholder quorum subsampling assumes fair sampling is correct. The algorithm implemented in the software package *R* (Alroy 2010a) subsamples within an interval until it reaches a user defined quota ( $\approx$ quorum or  $q$ ). The value represents the area under a frequency distribution curve of observed taxa (Alroy 2010a). The frequencies of taxa are analogous to the shares of those taxa in a company. The quorum is reached when a certain number of shares are represented. When taxa are first sampled, their share is credited, i.e., their frequency is added to the total amount of the distribution curve being sampled. Sampling the same taxa over and over does not affect the end results because their share is only credited the first time they are sampled. Subsampling continues until a fraction of frequencies (shares) are represented by at least one taxon. It is impossible to know the true frequencies of fossilised taxa because it is unlikely that one will ever sample every taxon known to have existed at any one point in time (Alroy 2010a). By assuming there were truly 10 taxa with a frequency of 0.1 each, instead of the reality of 100 and 0.01 each, one would overestimated the frequencies of those 10 taxa, and would draw only 2 on average when one should have drawn 20.

To counter this over estimation of frequencies, SQS calculates the “coverage” ( $u$ ) of taxa in a time interval i.e. the proportion of the whole frequency distribution represented by taxa found so far (Colwell & Coddington 1994; Good 1953). The algorithm is then free to continue subsampling until it reaches the level of coverage specified ( $q$ ). I have decided to use SQS over Rarefaction because the latter method can be very harsh if the lowest number of occurrences is a very small number e.g. one interval contains two taxa, the rest over 100, so each stage is rarefied to two taxa, flattening out the diversity curve.

One other resampling method for calculating diversity was considered that is a variant of the traditional counting of genera. It is worth special mention for the way in which it accounts for the uncertainty in stratigraphic ranges (Lloyd *et al.* 2011). Unlike traditional counting methods that fixate on counting taxa from interval to interval in specific ways ( $N_b$ ,  $N_t$ ,  $N_r$ , SIB, first differences), the method of Lloyd *et al.* 2011 focuses on the uncertainty of the stratigraphic ranges themselves. For example, one taxon may be found at an outcrop dated precisely to the Maastrichtian stage while another is dated broadly to the Cretaceous. Either or both of these dates could be correct e.g. one taxon was very restricted, the other wide ranging or of a higher taxonomic rank. The more conservative approach is to assume some level of bias in these ranges e.g. the more restricted taxon has simply not been looked for and is actually abundant across the Cretaceous, or the long ranged taxon was dated before more sophisticated dating techniques were available.

Traditionally this uncertainty has been dealt with by removing taxa that do not meet some *a-priori* assumption on range stability e.g. single stages accurate or family ranges accurate. The method of Lloyd *et al.* 2011 (Lloyd *et al.* 2011) assumes that singleton stratigraphic ranges are more accurate than longer ranging taxa. Rather than removing long ranging taxa; each was assigned a sub-range (single stage) of the total

observed range selected and the diversity curve calculated (SIB). This process is repeated 1000 times to produced 95% confidence limits on the diversity in each stage (Lloyd *et al.* 2011). The choice of a single stage is somewhat arbitrary but usually born from a detailed knowledge of the stratigraphic resolution of the group.

The method was designed around coccolithophores (unicellular eukaryotic phytoplankton with calcareous plates) which have one of the best fossil records of any organism (finely divided across the Phanerozoic). Singleton dates are common in coccolithophores. Longer ranging dates are likely to be an artifact of taxonomic under splitting than the coccolithophore being stratigraphically long ranging. Bivalves also have an impressive fossil record, if not to the same stratigraphic resolution as coccolithophores. The size of sub-range length used herein are single Stages as in the original method (Lloyd *et al.* 2011); however, I point out that this is the most conservative approach and could be overly harsh as a criterion for bivalve genus ranges which are on average longer than single stages. Future work is needed to assess this method across a range of clades at various taxonomic ranks to see what effect longer sub-ranges have on interpreting stratigraphic uncertainty.

The last approach to diversity estimation explored the use of modeling techniques to remove the bias of sampling using proxies (number of formations or number of collections) through time (Smith & McGowan 2007; Lloyd 2011; Lloyd *et al.* 2011). The method assumes that true diversity is constant though time and that observed diversity is controlled by a user-specified sampling proxy. By modeling the level of diversity expected using the proxy, spikes in diversity that fall outside the predicted results can be assumed to be genuine excursions requiring some explanation other than sampling. Using the data downloaded from the Paleobiology Database (Kiessling *et al.* 2013), the number of formations and collections containing bivalves were projected through time.

Once the number of formations and collections in each bin are identified an algorithm (Smith & McGowan 2007; Lloyd 2011) applies several regression models (linear, logarithmic, exponential, hyperbolic, sigmoidal and polynomial) to the data. The model that fits the data best is chosen via the Akaike Information Criterion (Johnson & Omland 2004). After selecting the best model to fit the data, expected values of diversity are calculated and the residuals (observed – expected) plotted through time. Error bars are calculated using standard error and standard deviation. Medium term trends (if picked up in the data) are also calculated using Multivariate Adaptive Regression Splines (MARS) (Friedman 1991; Lloyd 2011). It is a linear regression technique that accounts for nonlinear properties in data by calculating “hinge” points along the data series, revealing medium term trends that may not be visible otherwise.

To provide some context for the changes in diversity through time, I have also used the Paleobiology Database to plot the changes in bivalve diet and suspected life habit across the Phanerozoic. The entries in the database include a number of life habit

codes: nektobenthic, epifaunal, intermediate level epifaunal, low level epifaunal, semi infaunal, shallow infaunal, infaunal, deep infaunal and boring. These codes have been applied by the person who entered the data into the database and normally coded directly from the original source paper. There is a level of subjectivity in this classification system as the database does not classify the differences between codes such as semi infaunal and low level epifaunal. I therefore grouped some of these codes together into the same categories: nektobenthic, epifaunal (epifaunal, intermediate level epifaunal, low level epifaunal), infaunal (semi infaunal, shallow infaunal and infaunal), deep infaunal and boring.

There is a bias in grouping the codes this way. The categories epifaunal and infaunal included many more entries than nektobenthic, deep infaunal and boring due to this grouping of codes, which may inflate the numbers of epifaunal and infaunal. To counter this bias I counted the genera using Period as my level of stratigraphic resolution. Specifically I found the genera in each period and counted the number of life habits that belonged to those genera. Some genera were given several life habit codes. Codes such as intermediate level epifaunal and low level epifaunal were considered the same category (epifaunal in this case). If the genera fell into more than one category, e.g. epifaunal and infaunal, they were counted for each category they fell into. The analyses in this section assumed that taxa falling into several categories truly occupied several life habits; however, this may not be totally correct. It is highly likely some of categories given will have been the result of the subjectivity of these codes resulting in misclassification. If the degree of misclassification is high there is then the chance of inflation in the counts of categories through time due to this possible bias. Only a full revaluation of the given codes for all bivalve entries in the Paleobiology Database get at the real scale of this possible bias. This would be an enormous task given the bivalve entries number over 26,000 in the Paleobiology Database. The results present in this section relating to counts of life habit should therefore be viewed as preliminary requiring further investigation.

I used the same procedure for the counts of diets through time; however, I was not able to combine some categories that could effectively mix, e.g., suspension feeder could refer to carnivory or herbivory depending on what part of the suspension the bivalve was feeding on. I therefore assumed that these categories of diet (carnivory and herbivory) represent specific specialisations unique to suspension feeders or deposit feeders. The number of genera defined into these specific categories was small and so unlikely to affect the overall result. The categories for diet were: deposit feeder, suspension feeder, chemosymbiotic, photosymbiotic, carnivore, herbivore and grazer.

Finally I plotted change in sea level with time as found in Haq & Schutter (2002) and Snedden & Liu (2010). In summary, three methods: traditional counts, shareholder quorum and sampling proxy modelling were compared to assess the level of diversity in bivalves and three subclades that make up bivalves (Protobranchia, Heteroconchia and

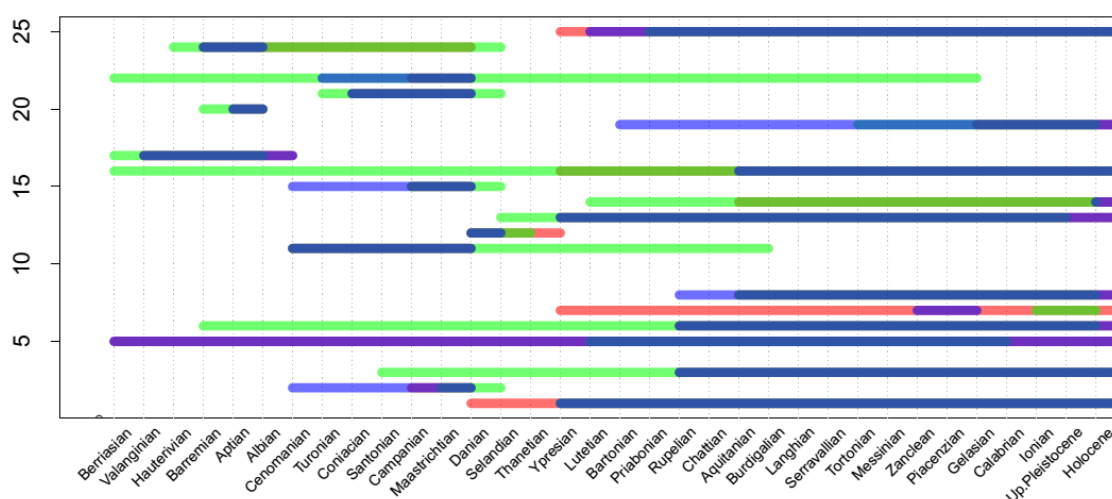


Fig.2.0 Overlap and/or differences in stratigraphic range (Early Cretaceous to the Present) for 25 genera shared by the Sepkoski (red), Paleobiology (green) and Treatise Database (blue). Darker colours represent overlap between the databases, lighter colours represent differences.

Pteriomorphia) across the Phanerozoic. As the latter two methods required data that comes as part of the Paleobiology Database download i.e. occurrence and sampling proxy information, these two subsampling methods were not applied to the Treatise stratigraphic data.

## 2.6 Results

### 2.7 Diversity Over the Phanerozoic: Updates in the Data

Of the original 2436 bivalve genera (Sepkoski Database), only 1582 are found in the current form of the Paleobiology Database. This implies the Paleobiology Database has replaced 854 genera from the original data. As the Paleobiology Database itself contains 2469 genera, an additional 33 genera have also been added that were not part of the Sepkoski data (Miller & Sepkoski 1988; Sepkoski 2002). Of those genera shared between the Sepkoski and Paleobiology Databases, 1424 genera (90%) have had their stratigraphic ranges modified by at least one stage. The modifications made by the Treatise to the Sepkoski data are different to the modifications made by the Paleobiology Database. The number of genera shared between the Sepkoski and Treatise is greater (2139). This implies that the Treatise database recognises the majority of genera in the Sepkoski Database as valid, only replacing 21 of the original genera.

The +800 genera difference between the Sepkoski and Paleobiology Database hides the high level of overlap in genera between the three databases e.g. the genus *Yabepecten* is found in the Sepkoski and Paleobiology Database but not the Treatise while the genus *Zoziella* is found in the Treatise and Sepkoski Databases but not the

Paleobiology Database. The Sepkoski Database contains 191 genera not contained in either of the other two databases, while the Paleobiology Database contains 342.

The revised Treatise is the largest contributor of unique genera, containing 1912 unique to that database. Of the unique genera in the revised Treatise, 814 are singletons (genera occupying a single stage), while the other 1082 are scattered throughout the rest of the Phanerozoic. The singleton genera are all stratigraphically located at the present, but make up less than half of the 1816 genera recorded as extant. Of those genera shared between the Sepkoski and Treatise databases (2139), 1447 (68%) have had their ranges modified by at least one stage.

There are a total of 1476 genera in common between all three databases. When the ranges of these genera are compared (Fig.2.0), all have modified by at least one stage in one of the databases, but there is a small number of genera (<100) whose range is equal across two of the three databases. Neither the Paleobiology nor Treatise databases agree on any of these genera. Where there is a difference between the three databases, in the Treatise the direction of change has been balanced (range extensions 678, range reductions: 713) while the Paleobiology Database has mostly extended ranges (844) than reduce them (484). The majority of these range extensions and reductions have been small for the Paleobiology Database: changes of +/-5 stages: total 860; while the Treatise has modified a broader set of ranges for an equivalent number of genera: changes of +/- 10 stages: total 846.

## **2.8 Traditional Counts (*Bivalves*)**

The counts of bivalves across the Phanerozoic based on the original Sepkoski Database, the Paleobiology Database and revised dates from the Treatise of Invertebrate Palaeontology are plotted through time in Figure 2.1. The clearest difference between the three curves is the increased number of genera contained within the revised Treatise (4596) compared to either the Paleobiology or Sepkoski Databases (2436 and 2469 respectively). By time interval, the largest difference between the curves resides at the present, where the Treatise data contains many more extant taxa than either database; however, the Treatise also contains approximately double the number of genera across the Palaeozoic and Mesozoic. There is a clear similarity between the curve of the Paleobiology Database and the Sepkoski Database, although there are some large differences in particular periods. Across the Permian and Triassic the Paleobiology Database has a greater number of genera compared to the Sepkoski Database while in the Cenozoic; this pattern is the other way round.

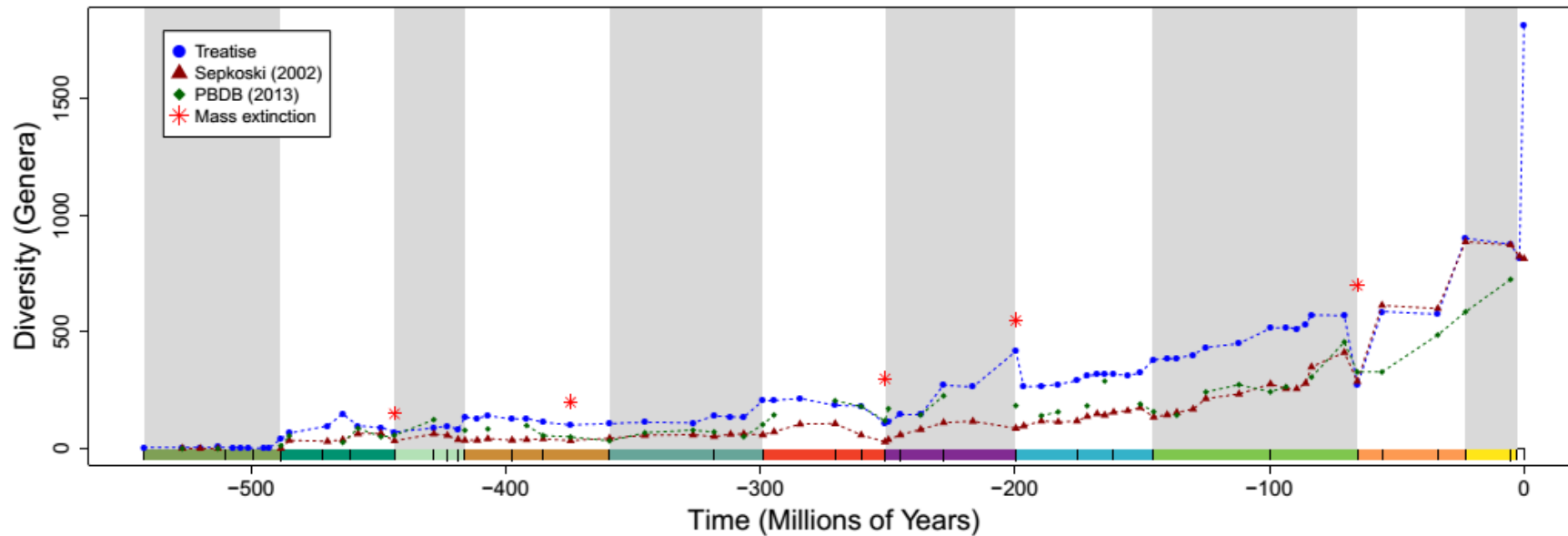


Figure.2.1. Diversity estimates (raw counts of genera in each bin) for Bivalvia across the Phanerozoic calculated from the data of Sepkoski Database, the Paleobiology Database, and revised dates from the Treatise of Invertebrate Palaeontology. Blue: Treatise; red: Sepkoski Database; green: Paleobiology Database. Diversity estimates are plotted using the base boundary values (millions of years ago) of the Sepkoski Database. Counts of genera are based on the traditional method of diversity estimation (add up the genera found in each stage). Red stars indicate timings of the top five mass extinction events: End Ordovician, Late Devonian, End Permian, End Triassic and End Cretaceous.

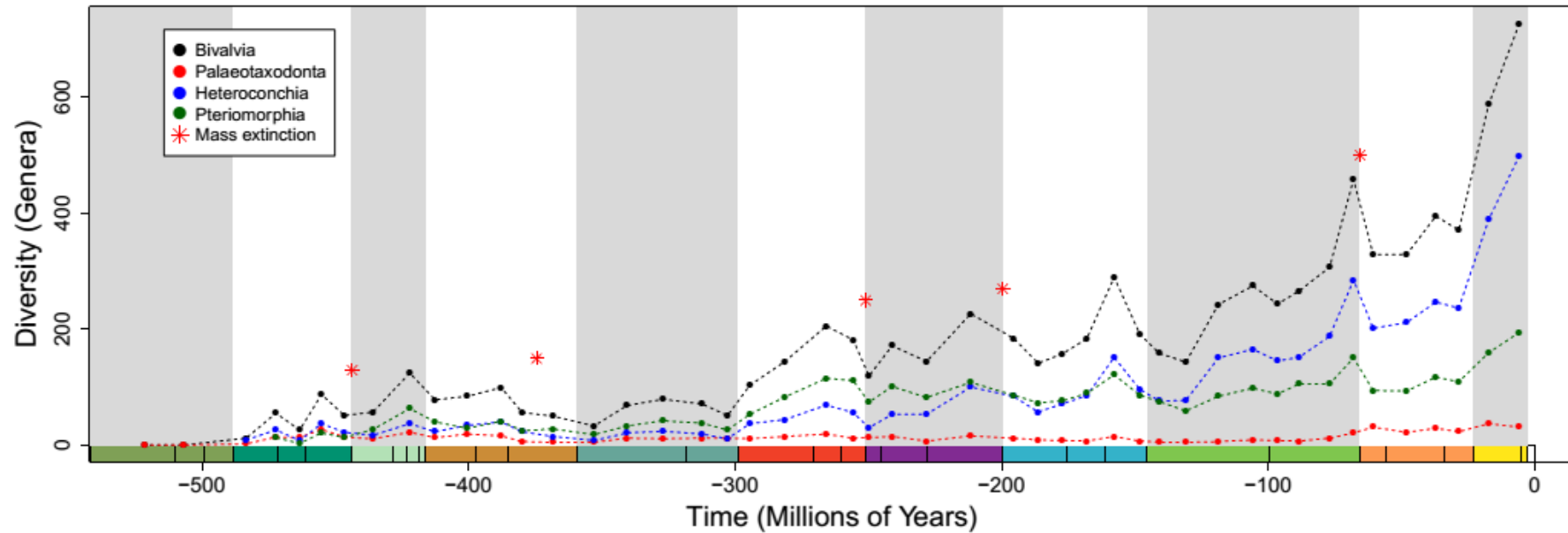


Figure.2.2. Diversity estimates (raw counts of genera in each bin) for Bivalvia, and its three major subclades (Protobranchia ( $\approx$  Palaeotaxodonta), Heteroconchia, Pteriomorphia), across the Phanerozoic calculated from the Paleobiology Database. Black: Bivalvia; red: Protobranchia ( $\approx$  Palaeotaxodonta); blue: Heteroconchia; and green: Pteriomorphia. Diversity estimates for each stage are centred on the midpoint of each stage. Counts of genera are based on traditional method of diversity estimation (add up the genera found in each stage). Red stars indicate timings of the top five mass extinction events: End Ordovician, Late Devonian, End Permian, End Triassic and End Cretaceous.

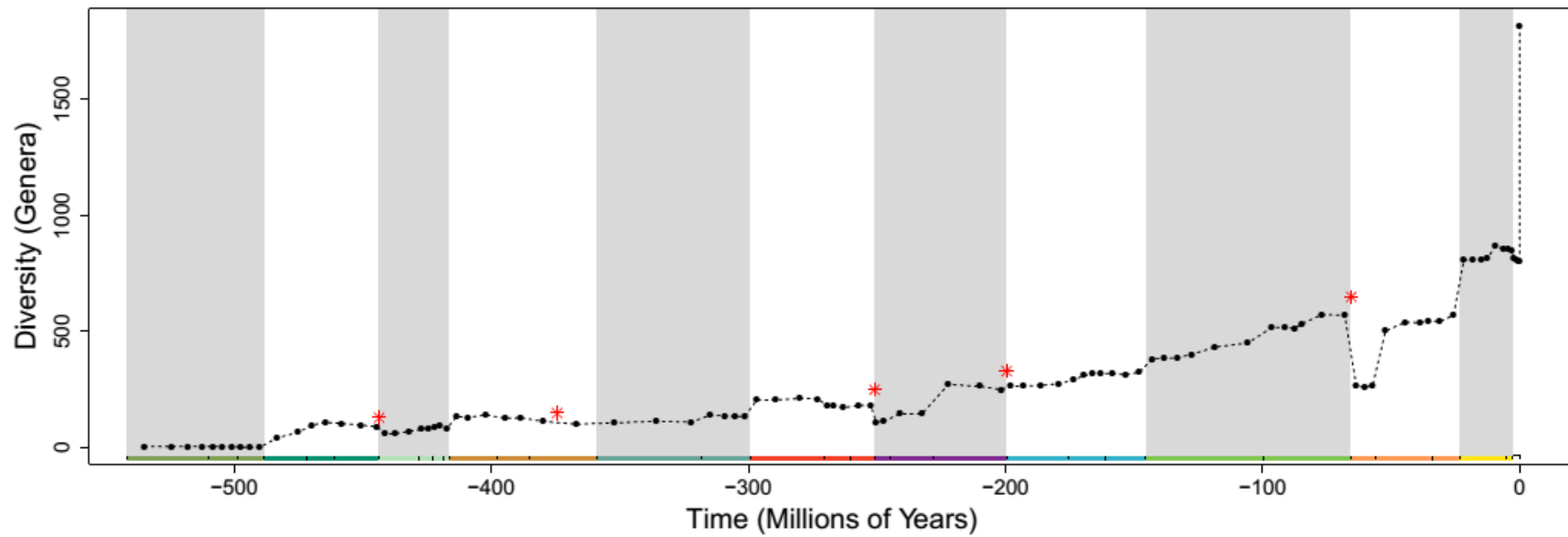


Figure.2.3. Diversity estimates (raw counts of genera in each bin) for Bivalves across the Phanerozoic calculated from the Treatise of Invertebrate Palaeontology (2013). (Black) Bivalvia. Diversity estimates for each stage are centred on the midpoint of each stage. Counts of genera are based on traditional method of diversity estimation (add up the genera found in each stage). Red stars indicate timings of the top five mass extinction events: End Ordovician, Late Devonian, End Permian, End Triassic and End Cretaceous.



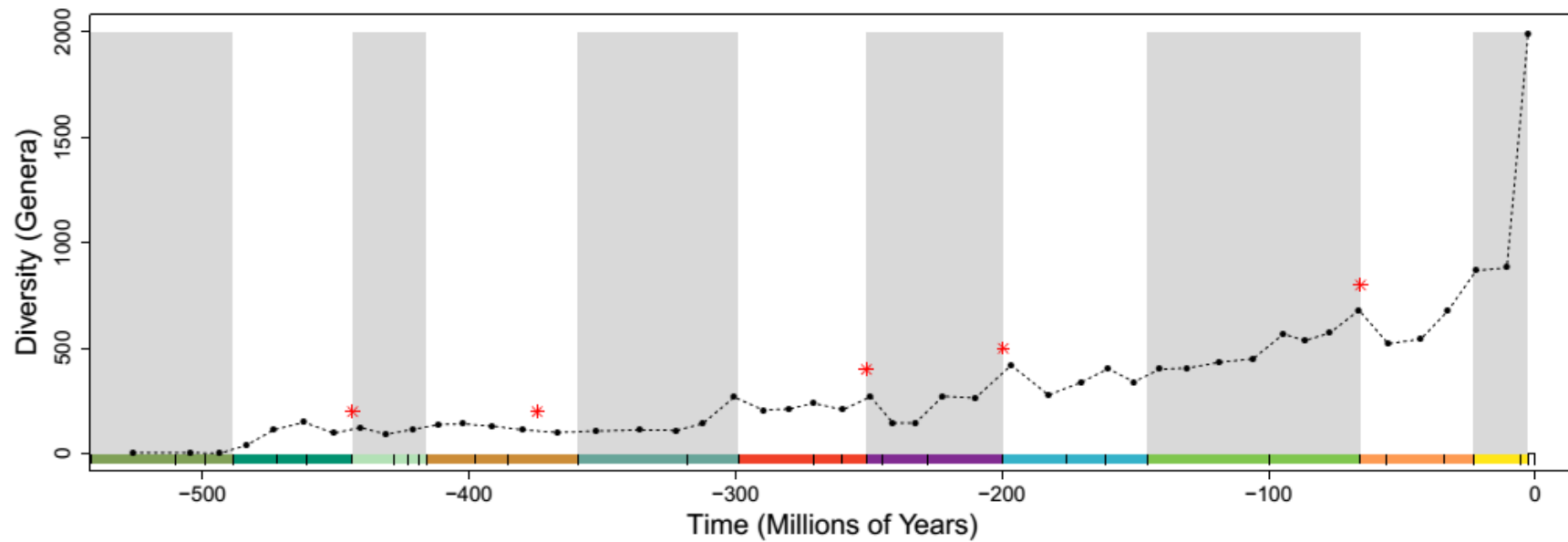


Figure.2.4. Diversity estimates (raw counts of genera in each bin) for Bivalves across the Phanerozoic calculated from the Treatise of Invertebrate Paleontology (*in prep*). Black: Bivalvia. Intervals have been split into roughly 48 bins of 10-11 million years each as in the Paleobiology Database. Counts of genera are based on traditional method of diversity estimation (add up the genera found in each stage). Red stars indicate timings of the top five mass extinction events: End Ordovician, Late Devonian, End Permian, End Triassic and End Cretaceous.

Despite additional data, the three curves show remarkable similarity. All three show the same classical increase in diversity through time. In the Palaeogene and Neogene the Treatise and Sepkoski curves show almost identical levels of diversity. Not only has the Treatise added more taxa to the Phanerozoic curve, 68% of the genera shared between the Treatise and Sepkoski Databases (1447 of 2139) have undergone some level of revision to their stratigraphic ranges. As for the Sepkoski Database and the Paleobiology Database, the number of genera shared is greater (1582); however, there has also been a greater revision of these taxa, as 90% of these genera (1424) have had their stratigraphic range revised in some way. The effect of these revisions on the shape of the curves seems to have been minimal compared to the effect of adding additional taxa.

The results of Figures 2.2-2.4 show diversity (traditional counts) of bivalves (black line) across the Phanerozoic. Figure 2.2 shows the diversity of bivalves based on the entries in the Paleobiology Database while Figure 2.3 shows the diversity based on the ranges compiled from the revision of the Treatise of Invertebrate Palaeontology for Bivalvia (*in prep*), supplied by Dr Joseph Carter (University of North Carolina). Both Figures agree on the general diversity trend of bivalves across the Phanerozoic. Bivalves appear in the Cambrian with only the five genera, growing in diversity to approximately 100-150 genera through the Ordovician, peaking in the Middle Ordovician.

Both curves show a drop in diversity at the end of the Ordovician, with a recovery through the Silurian. In the Devonian, Figure 2.2 shows a gradual drop from approximately 100 to 50 genera while in Figure 2.3 the drop in diversity is very shallow, with diversity remaining within the range of 100-150 up to the end of the Carboniferous. Figure 2.2 shows an increase across the Carboniferous from approximately 50 to 70-80 genera. Both figures show an increase in generic diversity through the Permian, up to approximately 200 genera. Each curve then suffers a drop in diversity at the End of the Permian and both show a recovery through the Triassic back to pre-Permian extinction levels. Figure 2.3 does not show a drop at the end Triassic extinction, which is picked up by the Paleobiology Database derived ranges in Figure 2.2. Through the Jurassic to the End of the Cretaceous, the Treatise data of Figure 2.3 shows a gradual increase in diversity to over 500 genera.

Figure 2.2 shows a slightly different trajectory, with a drop in diversity through the End Triassic mass extinction, similar in scope to the end Permian Mass extinction, followed by another recovery, peaking at approximately 350 genera in the Late Jurassic. There is another drop between the Late Jurassic and Early Cretaceous in Figure 2.2, with another rise up to the End Cretaceous with over 400 hundred genera. Figures 2.2 and 2.3 agree that the End Cretaceous extinction was the largest of its kind for bivalves, with diversity dropping 100 genera in Figure 2.2, and 250 genera in Figure 2.3. Finally both

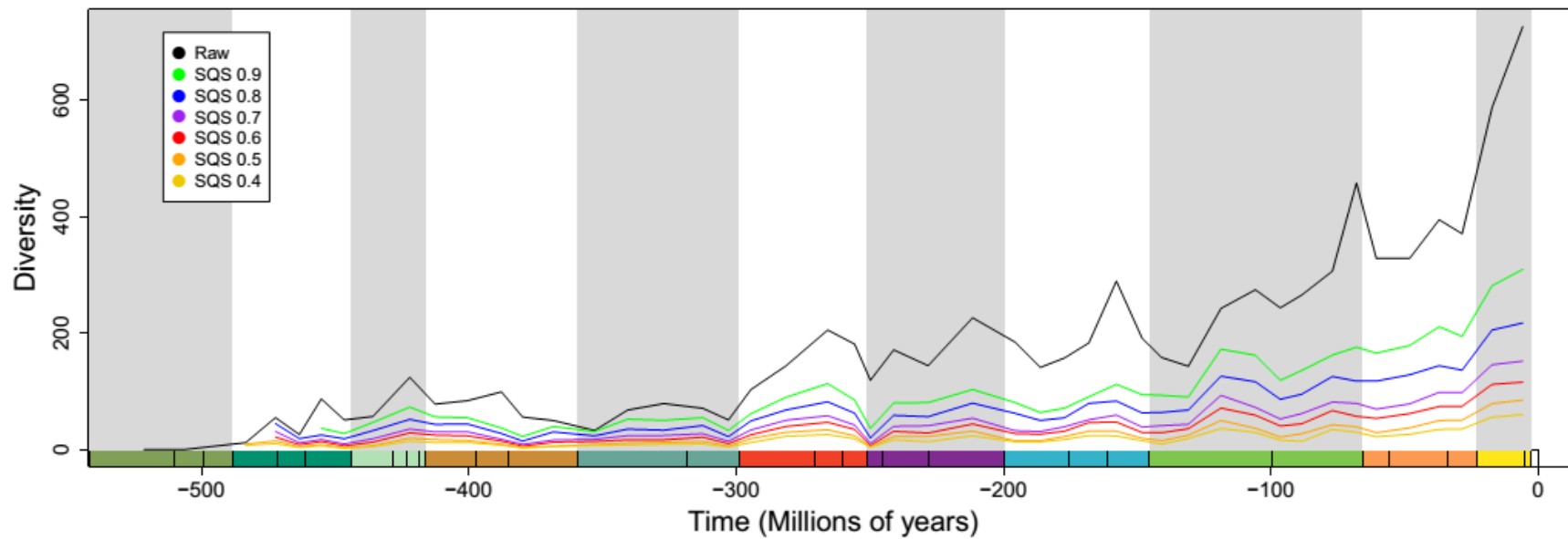


Figure.2.5. Diversity estimates (number of genera) for Bivalvia across the Phanerozoic as calculated from the Paleobiology Database online facilities. Duration of each bin: ~10-11 million years. Black: counts of genera based on PBDB default settings without sampling standardisation (Traditional counts method). Shareholder quorum subsampling levels: green: 0.9; blue: 0.8; purple: 0.7; red: 0.6; orange: 0.5; yellow: 0.4. Each sub-sampling method used the maximum number of replications available via the website (100 replicates).

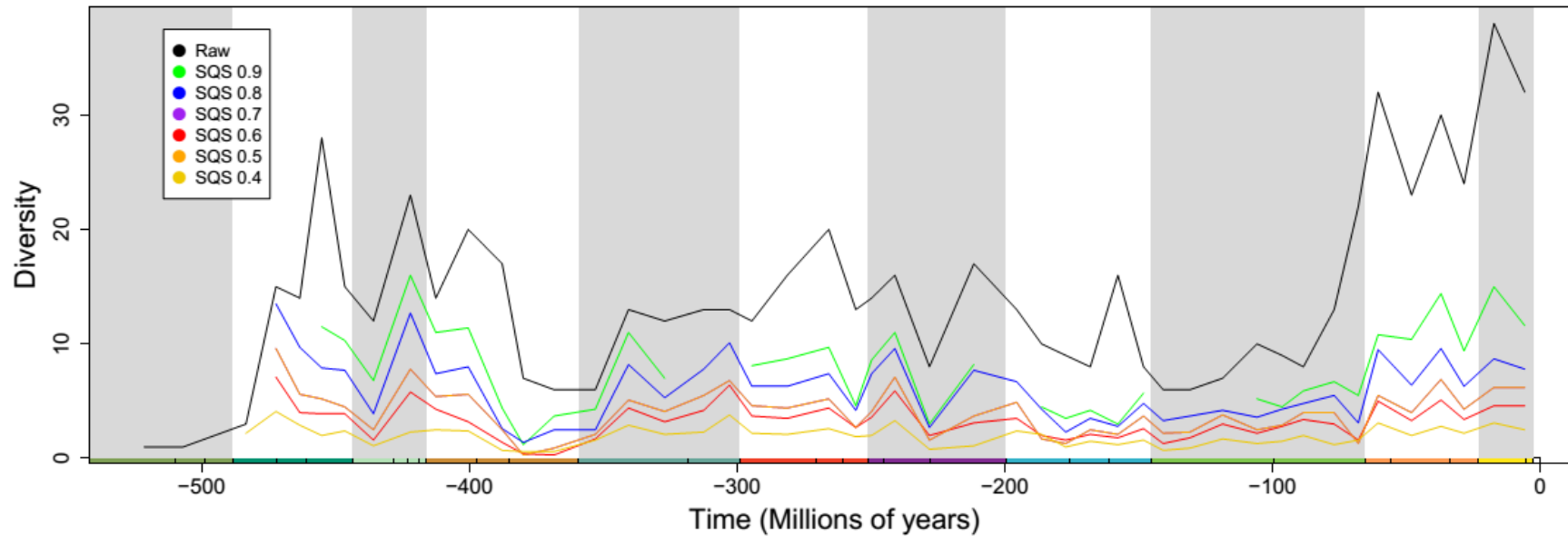


Figure.2.6. Diversity estimates (number of genera) for Protobranchia ( $\approx$  Palaeotaxodonta) across the Phanerozoic as calculated from the Paleobiology Database online facilities. Duration of each bin:  $\sim$ 10-11 million years. Black: counts of genera based on PBDB default settings without sampling standardisation (Traditional counts method). Shareholder quorum subsampling levels: green: 0.9; blue: 0.8; purple: 0.7; red: 0.6; orange: 0.5; yellow: 0.4. Each sub-sampling method used the maximum number of replications available via the website (100 replicates).

curves agree on a recovery through the Palaeogene and Neogene to their maximum level of diversity at present day.

The results of Figure 2.2 agree with generic diversity curves of past studies (Miller & Sepkoski 1988; Jablonski et al. 2003; Ros & De Renzi 2005; Kidwell 2005; Dmitriev & Nevesskaja 2006; Nevesskaya 2008; Vinarskia et al. 2011). This is unsurprising as each of these past studies have used the Paleobiology Database data to construct curves. The main difference between the two curves is the greater genera per interval recorded in the revised Treatise (Fig.2.3) compared to the Paleobiology Database (Fig.2.2). From the Devonian period onwards, Figure 2.3 does not vary across periods, resulting in a much flatter trajectory than Figure 2.2. Only the End Permian and End Cretaceous mass extinctions cause any noticeable drops in diversity.

This difference between the two curves is partly an artefact of the number of time intervals used in each Figure. Figure 2.4 shows the diversity of bivalves using the revised Treatise dates once converted into the same 10-11 million year intervals as Figure 2.2. Some of the flatness is removed between Figures 2.3 and 2.4 but not all of it. The same general pattern of relatively low, constant Palaeozoic diversity followed by a post-Palaeozoic rise is still evident. The main differences between the trajectories of the diversity across the Phanerozoic between the Paleobiology Database (Fig.2.2) and Treatise (Fig.2.3-2.4) is found in the Late Silurian, Middle Devonian, Late Permian and Late Jurassic period. During the Late Silurian, the Paleobiology Database picks up a prominent peak in diversity prior to a drop across the Pridoli-Lockhovian boundary (416 million years ago). This peak is greater than the peak in the Ordovician (Fig.2.2). Figures 2.3 and 2.4 do not pick up this Silurian peak, instead diversity peaks earlier, i.e., during the Middle Ordovician, timed alongside the great Ordovician biodiversification event (Cope 2004).

The Middle to Late Devonian decline of bivalves shown in Figure 2.2, missing in Figure 2.3 is also missing in Figure 2.4. Instead, Figure 2.4 shows a roughly stable amount of diversity up to the end of the Carboniferous where diversity increases from 100-150 to approximately 250 genera. Through the Permian, diversity fluctuates around 250 until the End Permian extinction. In Figure 2.4 the extinction occurs mostly in the Early Triassic. The pattern has been identified in other diversity patterns (Miller & Sepkoski 1988; Nevesskaya 2008). Across the Triassic, Figure 2.4 shows an increase in diversity which was evident in Figures 2.2 and 2.3. Unlike Figure 2.3 which shows no drop in diversity across the End Triassic mass extinction, Figure 2.4 does pick up a decrease in diversity of approximately 100 genera. From the Jurassic to the present, Figure 2.4 then follows a similar trajectory as the data in Figure 2.3.

## **2.9 Traditional Counts (Subclades)**

Figure 2.2 shows the traditional counts for the three subclades: Protobranchia, Heteroconchia and Pteriomorphia. In modern oceans, Protobranchia make up the smallest proportion of diversity (Sharma *et al.* 2013), which is the general trend across the Phanerozoic (Fig.2.2 red line). Heteroconchia and Pteriomorphia are found in greater abundance through the entire Phanerozoic but in unequal amounts, taking turns as the most abundant clade (Fig.2.2 blue and green lines). During the Ordovician period, Heteroconchia are the most abundant clade; however, after the End Ordovician extinction, Pteriomorphia rise to dominance over the Heteroconchia. Through the Devonian period all three clades show a decline in diversity, with Heteroconchia and Pteriomorphia switching roles as most dominant clade across the period. Pteriomorphia regain dominance at the end of the Devonian period, with diversity remaining stable in the two infraclasses across the Carboniferous. In the Permian, Pteriomorphia and Heteroconchia begin to rise in diversity while Protobranchia remain in low abundance up to the Middle-Late Permian (Guadalupian).

These two infraclasses suffer extinctions across the boundary but both recover most of their lost diversity by the Late Triassic. Across the End Triassic boundary both clades show a gradual decline into the Early Jurassic, but from the Middle Jurassic to the End Jurassic diversity rose again.. At this point Heteroconchia retake their place as the most dominant clade. From this point onward, both the Heteroconchia and Pteriomorphia follow diversity trajectories similar in shape to the curve for bivalves (Fig.2.2 black line); however, Pteriomorphia do not regain their title of most generically abundant clade.

## **2.10 Shareholder Quorum Subsampling (Bivalvia)**

The results of the shareholder quorum subsampling results for bivalves across the Phanerozoic are found in Figure 2.5. The colours green, blue, purple, red, orange, and yellow refer to quorum levels of 0.9 to 0.4, representing 90-40% diversity sampling. The results were obtained using the online tools of the Paleobiology Database and plotted using R. For easy comparison, the Figure also shows the results using traditional counts, just as in Figure 2.2. Whichever quorum level is used, the trend of diversity increasing across the Phanerozoic remains true. The steep exponential growth curve as suggested by the traditional counts is gone, replaced by a more shallow increase through time (Spearman Rank correlation  $\rho = 0.905$ , p-value < 2.2e-16).

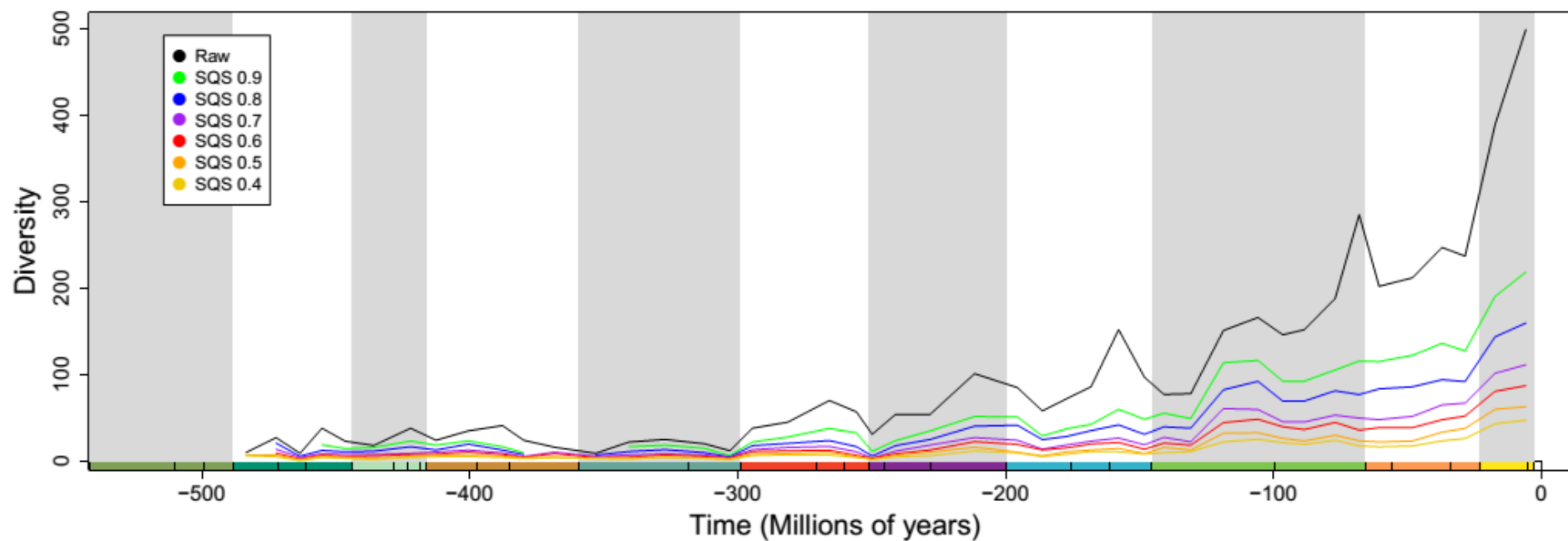


Figure.2.7. Diversity estimates (number of genera) for Heteroconchia across the Phanerozoic as calculated from the Paleobiology Database online facilities. Duration of each bin: ~10-11 million years. Black: counts of genera based on PBDB default settings without sampling standardisation (Traditional counts method). Shareholder quorum subsampling levels: green: 0.9; blue: 0.8; purple: 0.7; red: 0.6; orange: 0.5; yellow: 0.4. Each sub-sampling method used the maximum number of replications available via the website (100 replications).

Many of the rises and falls across the different periods disappear as the quorum level is reduced. With a quorum level of 0.9, the peak at the End Jurassic is very much reduced (from 250 genera to 100). The End Cretaceous peak and proceeding extinction, found in both the traditional counts of the Paleobiology Database and Treatise datasets, disappears completely. It is replaced by a gradual rise across this boundary. The End Permian mass extinction remains under each quorum level, suggesting that this decrease in diversity is robust to resampling procedures.

### **2.11 Shareholder Quorum Subsampling (Subclades)**

The shareholder quorum subsampling curves for the individual subclades are found in Figures 2.6-8. The diversity of Protobranchia based on traditional counts and on the SQS values for Protobranchia are found in Figure 2.6. The diversity of this clade across the Phanerozoic is uneven (Fig.2.6. black line). The most likely reason for this fluctuation is the low number of entries found in the Paleobiology Database. According to Figure 2.6, diversity in the Protobranchia rises quickly through the Ordovician, reaching a peak of over 25 genera. This level of diversity is not reached again until the End Cretaceous boundary, where diversity rises across the extinction boundary to over 30 genera from roughly 10 genera in the Middle Cretaceous.

Between these two time periods, diversity fluctuates between 10 and 20 genera, but is steadily dropping. Despite the fluctuations some aspects of the bivalve diversity curve (Fig.2.2) are still present. There is an End Silurian peak, followed by a Devonian drop and Carboniferous-Permian gradual increase in diversity. The results of the shareholder quorum subsampling also show fluctuations between the periods. There are a few gaps where the diversity is not high enough to reach the specified quorum level. Where shareholder quorum levels can be reached, the same pattern as the traditional counts can be picked out.

Figure 2.7 shows the traditional counts in diversity of Heteroconchia across the Phanerozoic. The most striking aspect of this curve is its similarity to the curve for bivalves as a whole. This suggests that for the Paleobiology Database the majority of entries on bivalves are contributed by Heteroconchia. The Palaeozoic part of the curve in Figure 2.7 is mostly flat until the Permian, where diversity begins to pick up. The End Permian mass



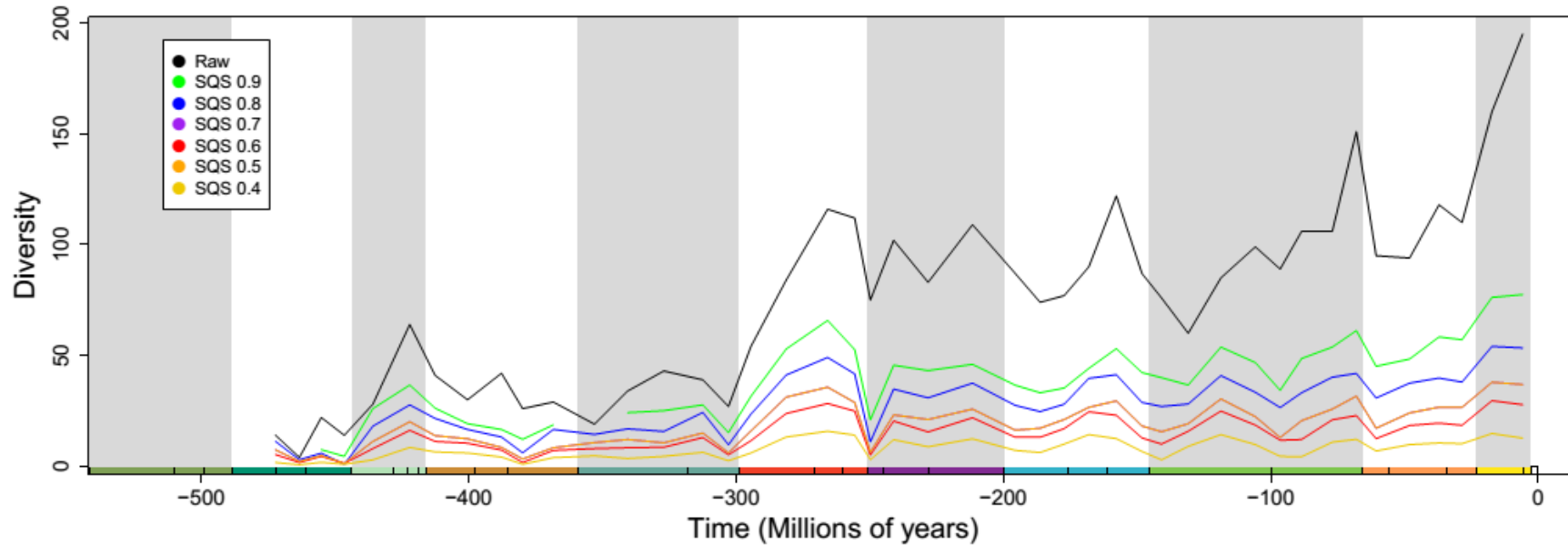


Figure.2.8. Diversity estimates (number of genera) for Pteriomorphia across the Phanerozoic as calculated from the Paleobiology Database online facilities. Duration of each bin: ~10-11 million years. Black: counts of genera based on PBDB default settings without sampling standardisation (Traditional counts method). Shareholder quorum subsampling levels: green: 0.9; blue: 0.8; purple: 0.7; red: 0.6; orange: 0.5; yellow: 0.4. Each sub-sampling method used the maximum number of replications available via the website (100 replicates).

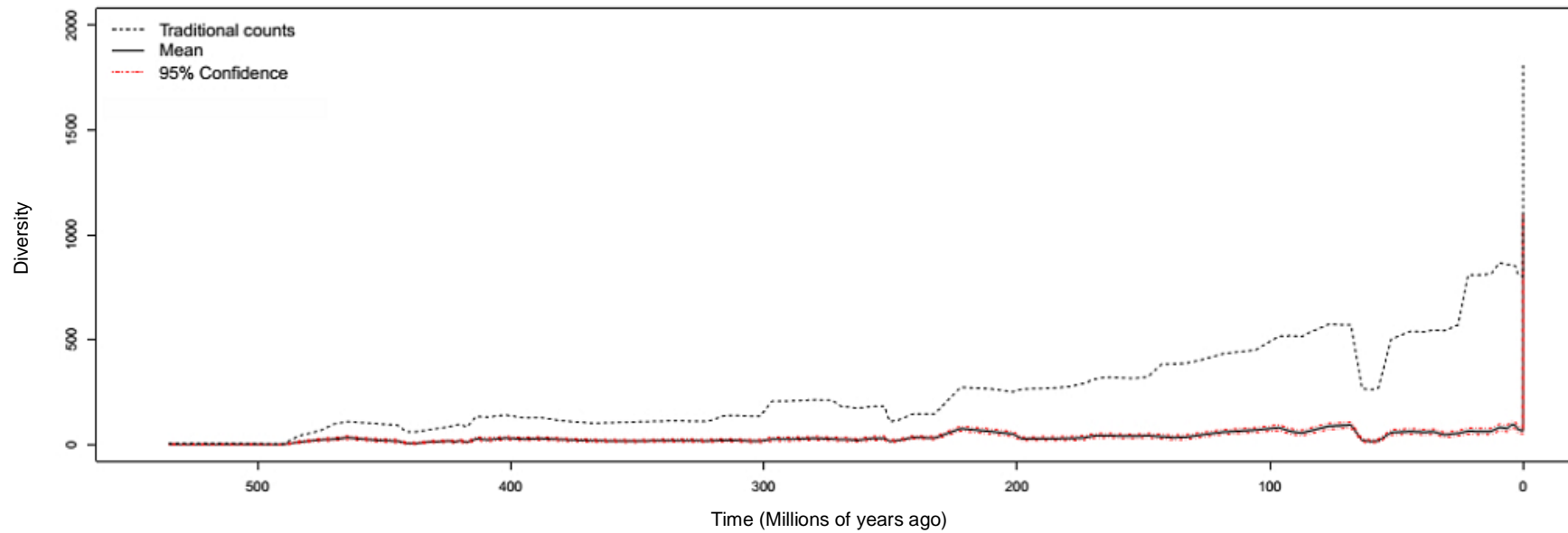


Figure.2.9. Diversity of bivalves using the Treatise ranges and the range uncertainty method. Traditional counts are highlighted in the dashed black line. Solid black lines are the mean of 1000 bootstrap replicates of diversity (traditional counting) after random range reassignment. The 95% confidence limits around the mean.

extinction did not affect Heteroconchia too severely. After the event diversity continues to increase up to the present. Like the shareholder quorum subsampling results of bivalves (Fig.2.5), once subsampling is used, the end Jurassic peak is gone, and there is no peak or fall across the End Cretaceous boundary.

The results for Pteriomorphia diversity using shareholder quorum subsampling are found in Figure 2.8. This clade shows the greatest increase in diversity across the Permian, rising from under 50 genera at its base to over 100 genera by the End Permian mass extinction. Prior to this, traditional counts show a peak of +50 genera in the End Silurian, with a fall through the Devonian. After the End Permian extinction, the diversity of Pteriomorphia fluctuates until the present between 70 (End Jurassic/Early Cretaceous boundary) and up to 200 genera (present). The shareholder quorum diversity curves suggest similar fluctuations as the traditional counts but smaller in magnitude.

Unlike the exponential increase conveyed by the traditional counts, shareholder quorum levels suggest more gradual increase, mostly due to a drop in diversity in the post Palaeozoic diversity of bivalves. Again, whatever the quorum level, the End Permian mass extinction makes a significant impact on Pteriomorphia diversity, whereas the End Cretaceous diversity drops in significance as the loss of 50 genera becomes a loss of 10 genera (Fig.2.8).

### **2.12 Range Uncertainty Method (Bivalves)**

The method of assigning shorter durations at random to those individuals with longer ranges (Lloyd *et al.* 2011; Lloyd *et al.* 2012) was applied to bivalves across the Phanerozoic using the Treatise dataset to see what effect this method has on the diversity. It is immediately obvious that this method removes much of the diversity signal seen in the Figure 2.3. Diversity becomes stable across the entire Phanerozoic up to the present, where diversity increases dramatically. Prior to this, diversity reaches its peak in the Middle Triassic. Figure 2.9 agrees with Figure 2.3 that the diversity decline in the End Cretaceous is the most significant drop across the Phanerozoic.

### **2.13 Sample Bias Modelling Diversity (Bivalves)**

Figures 2.10 shows the two sampling proxies: number of collections and number of bivalve bearing formations through time are based on the Paleobiology Database data. For comparison, the traditional counts of generic diversity are plotted alongside. Of the two curves, the number of collections through time shows the greatest similarity to the upward trajectory of diversity through time. A Spearman Rank correlation test between diversity and number of collections is highly significant (p-value = 5.307e-09) and positive ( $\rho = 0.731$ ). This is not surprising given that the Paleobiology Database is collection based. Several of the diversity spikes in the traditional counts curve (Fig.2.10 black line) are reflected by large spikes in the number of collections at those times including: the End

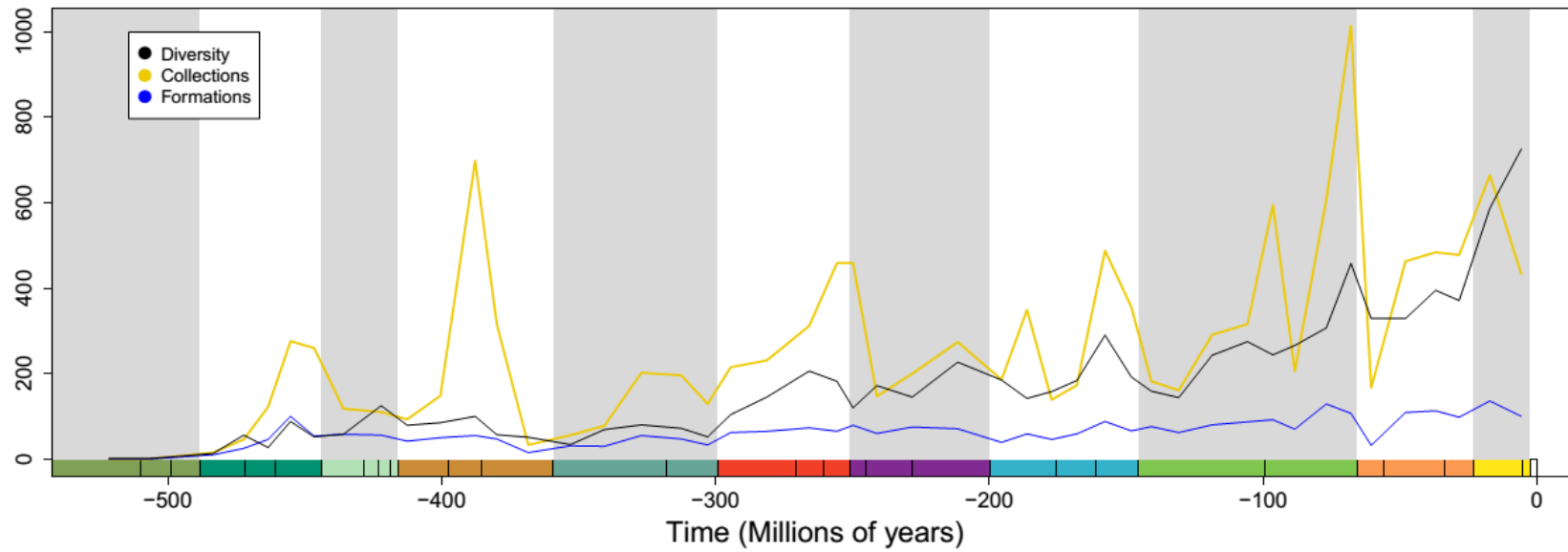


Figure.2.10. Number of formations and number of collections through the Phanerozoic. Generic diversity (traditional counts) highlighted in black. Data downloaded from the Paleobiology Database. Yellow line: number of collections through time. Blue line: number of formations through time.

Ordovician, the Late Devonian, the End Permian, the Late Triassic and the End Cretaceous. The diversity highs in those intervals coincide with the five mass extinction events.

Counting the number of bivalve bearing formations through time shows an entirely different pattern. Unlike the diversity and collections record, which increase through time, the number of formations remains almost constant, showing little perturbation, even at four of the five mass extinction events. The exception is a drop in formations at the End Cretaceous boundary. In this instance the number of bivalve forming formations drops by half. This flat trajectory suggests that the number of formations may be less of an issue in affecting the diversity of bivalves through time. When diversity and number of formations were tested for correlation using a Spearman Rank test, the results suggest again a highly significant and positive correlation ( $\rho = 0.806$ ,  $p\text{-value} = 8.559\text{e-}12$ ).

Each sampling proxy was used to model the diversity of bivalves through the Phanerozoic (Fig.2.11-2.12). In Figure 2.11 is the diversity expected based on the number of collections through time (Fig.2.11 red line) vs. the traditional counts diversity (Fig.2.11 blue line). The red dashed lines of this Figure (and Fig.2.12) show the diversity expected to be due to the sampling proxy. The outer dashed line represents the 95% confidence limits, and therefore any diversity estimates falling within these limits can be seen as to be entirely the result of the sampling proxy. For both sampling proxies, the model chosen by the script (Lloyd 2011) was the polynomial model:

$$y \sim x + I(x^2) + I(x^3) + I(x^4)$$

This model was chosen because it provided the best fit to the data given its AICc scores: 423.631 and 389.262 for number of collections and formations respectively. For much of the Palaeozoic and Mesozoic, there looks to be a close relationship between the number of collections and the traditional levels of diversity.

There are specific intervals where observed diversity and expected diversity do not overlap in Figure 2.11. At four of the five mass extinctions events i.e. the End Ordovician, Late Devonian, End Permian and End Cretaceous, the traditional diversity counts are found outside the lower bound of the 95% confidence limits of the expected diversity. This suggests that during these events there is less diversity present than expected given the number of collections at those times. The End Triassic, showed the opposite pattern, containing a greater diversity than expected given the sampling proxy and model.

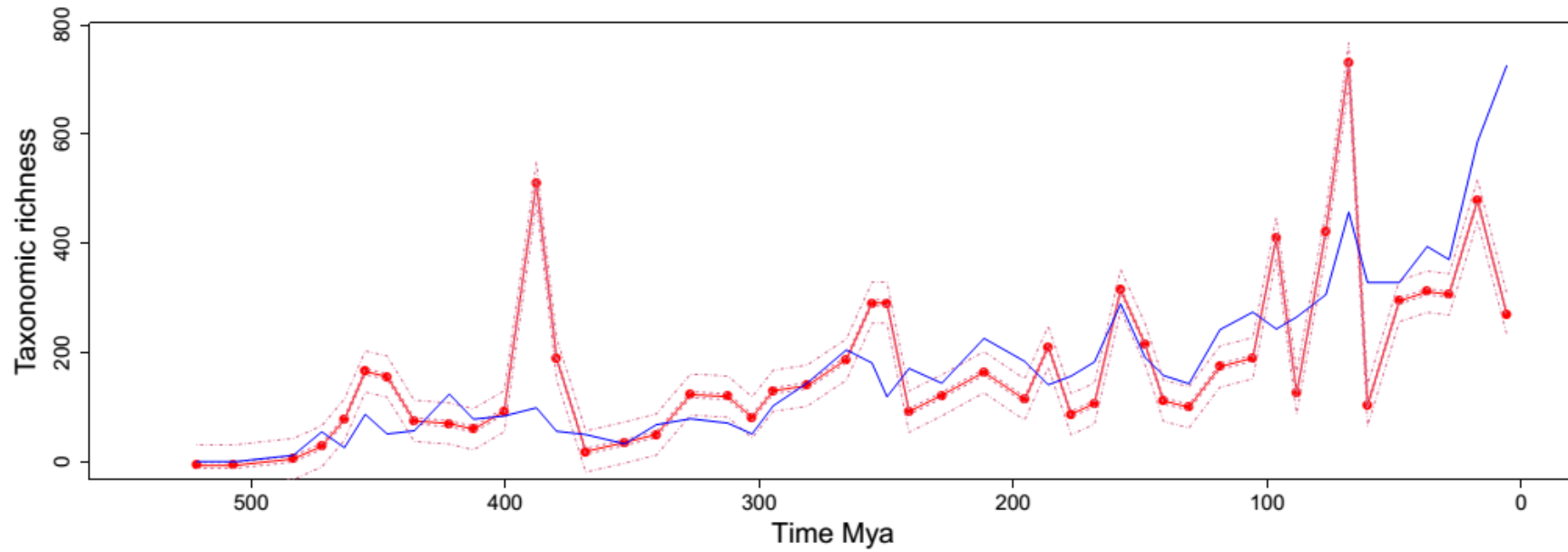


Figure.2.11. Generic diversity (traditional counts) vs. the diversity expected given the number of collections through the Phanerozoic, using number of collections as a sampling bias proxy. Blue line is the generic diversity based on traditional counts. Red solid line is the expected mean diversity. Inner dashed red lines indicate the confidence limits based on standard deviation while the outer red dashed line is the confidence limits of expected diversity levels based on standard error.

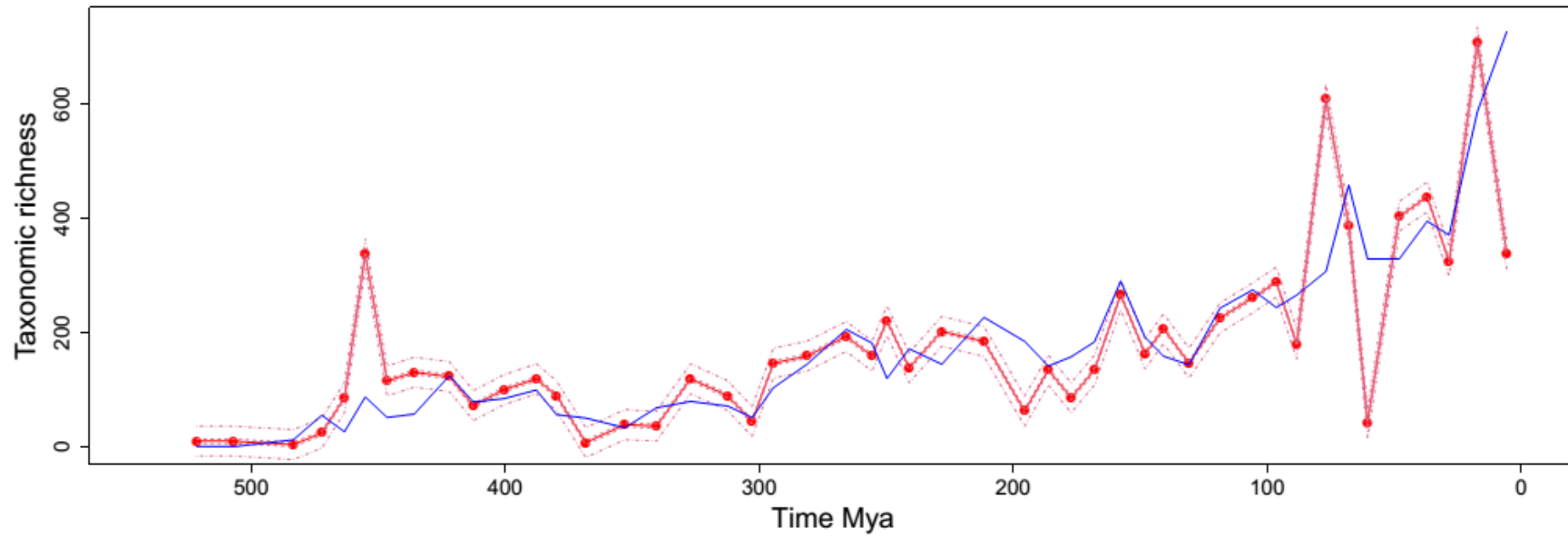


Figure.2.12. Generic diversity (traditional counts) vs. the diversity expected given the number of formations through the Phanerozoic, using number of collections as a sampling bias proxy. Blue line is the generic diversity based on traditional counts. Red solid line is the expected mean diversity. Inner dashed red lines indicate the confidence limits based on standard deviation while the outer red dashed line is the confidence limits of expected diversity levels based on standard error.

## **2.14 Detrended Diversity (Bivalves)**

The detrended diversity (observed diversity – expected diversity) based on the number of collections or formations can be found in Figures 2.13 and 2.14. These results are effectively the same as those found in Figures 2.11 and 2.12 respectively, but plotted to give a clearer picture of the deviations in diversity from the model predicted diversity. As in Figure 2.11, Figure 2.13 shows that the majority of bivalve diversity, especially in the Palaeozoic, is what would be expected to be the result of sampling. There are five exceptions: the End Ordovician, End Silurian, Middle Carboniferous, Late Devonian and End Permian. During the extinction events (End Ordovician, Late Devonian and End Permian), there are far fewer genera than expected given the number of collections, shown residuals far below the lower bound of the 95% confidence limit of expected diversity. Of these three extinctions, the Late Devonian shows the greatest diversity drop.

The late Silurian shows a positive deviation, finding diversity to be much higher than expected given the number of collections while the Middle Carboniferous shows a level of diversity less than expected by the number of collections. The residuals of the Mesozoic and Cenozoic (except for the Cretaceous period) lie above the upper bound of the 95% confidence limits, suggesting greater diversity than expected by the model. The end Triassic mass extinction is picked up, but shows the smallest deviation away from the expected results. The results of the MARS results are also plotted to show up any midterm trends that may not be apparent. For the analyses using number of collections as the sampling proxy the MARS results found no midterm trends that deviate from the model until the last 100 million years. At this point the technique recognises one hinge point, where diversity rises from the expected model, growing until the present.

The detrended results based on the observed diversity – expected diversity using the number of formations is seen in Figure 2.14. There are some noticeable differences between the detrended results of Figure 2.13 and 2.14 that were not readily visible in Figures 2.11 and 2.12. In the Palaeozoic the greatest deviations from the model occur in the Ordovician-Silurian period. During these two periods, diversity is far less than expected to be the result of sampling. Of the two periods, the End Ordovician mass extinction event shows the greater deviation, and is the second largest of its kind through the Phanerozoic (Fig.2.14). Through the Devonian, Carboniferous and Permian, much of the diversity remains within the limits of the modelled diversity. There is a switch across the Devonian/Carboniferous boundary from diversity lower than expected by the model to levels higher than expected. There is then a reverse during the Middle Carboniferous from greater than expected diversity to less than expected.

Through the Permian, diversity begins less than expected by the model, but rises into the range of model expectations until the End Permian mass extinction. During this event there is a drop in diversity compared to the number of formations similar to that



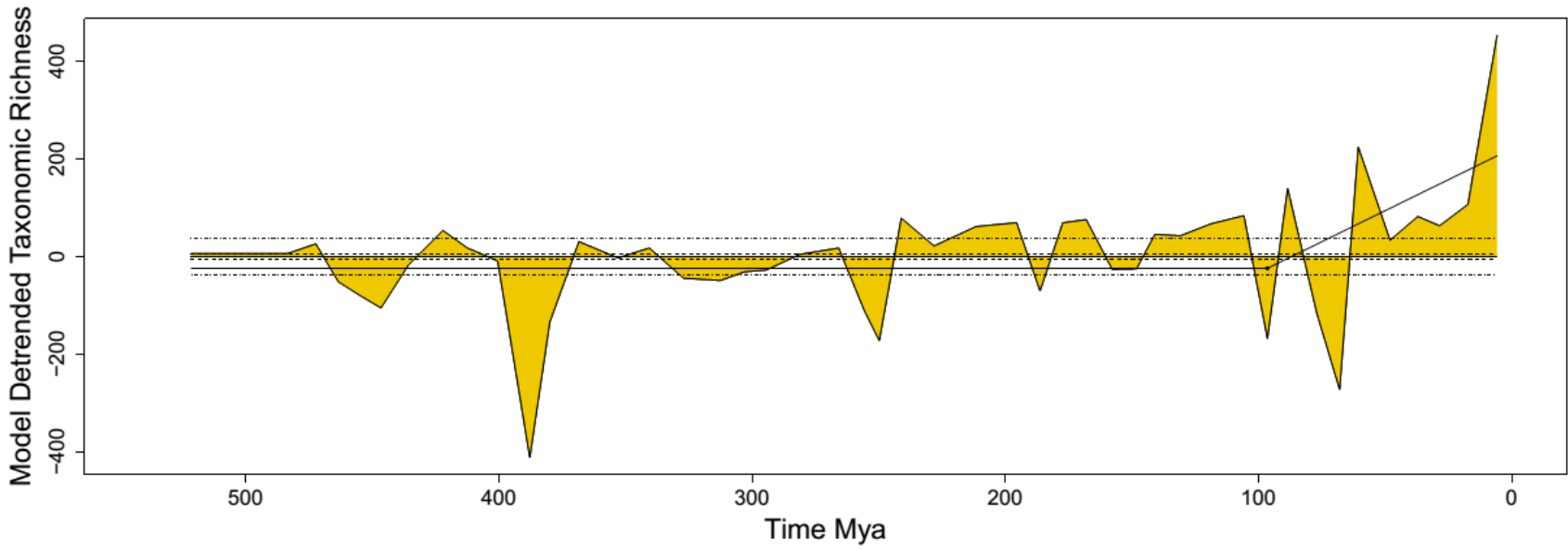


Figure.2.13. Detrended generic diversity (traditional counts) through the Phanerozoic, using number of collections as a sampling bias proxy. Dashed lines indicate diversity levels assumed to be the result of sampling bias. Dashed lines indicate 95% limits based on the standard error (out dashes) and standard deviation (inner dashes).

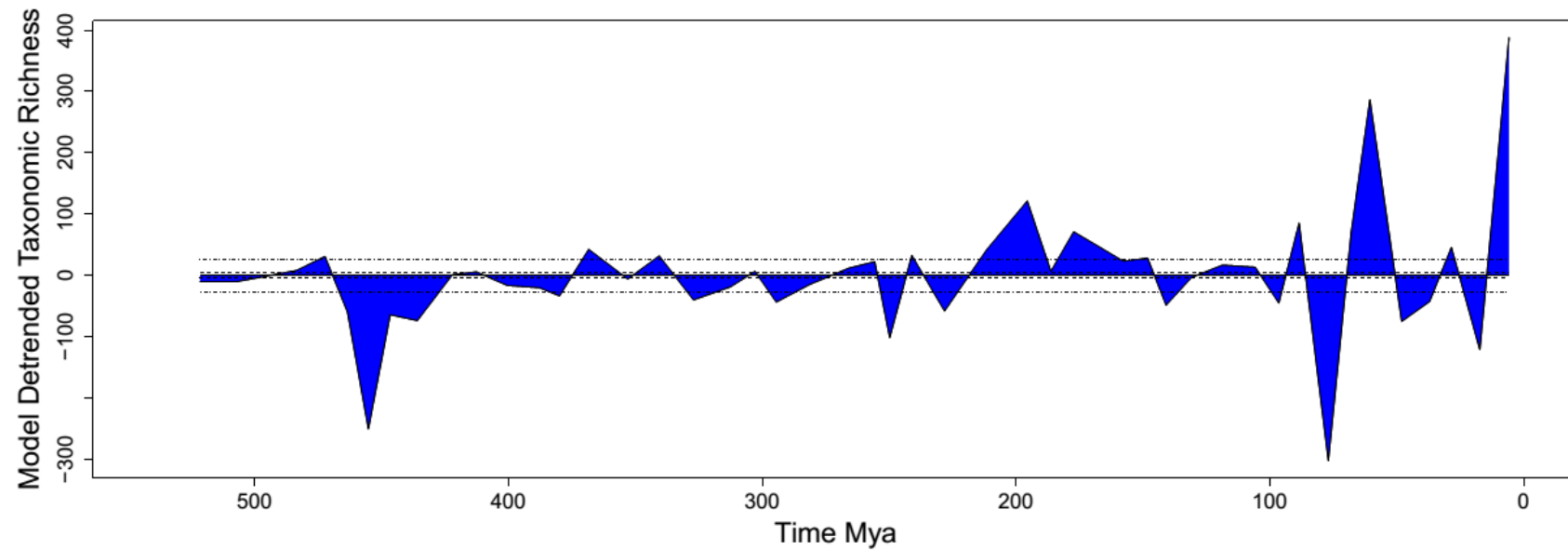


Figure.2.14. Detrended generic diversity (traditional counts) through the Phanerozoic, using number of formations as a sampling bias proxy. Dashed lines indicate diversity levels assumed to be the result of sampling bias. Dashed lines indicate 95% limits based on the standard error (out dashes) and standard deviation (inner dashes).

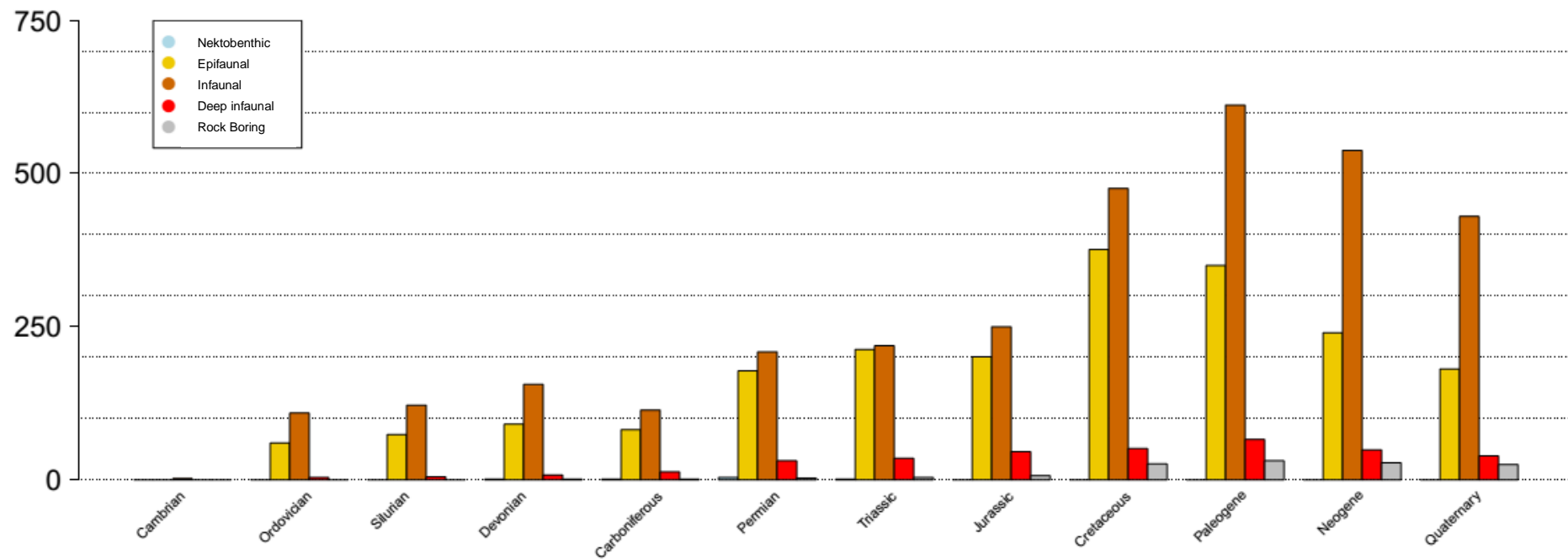


Figure.2.15 Counts of the life habits occupied by bivalves across the Phanerozoic as recorded in the Paleobiology Database, split by Period. The colours represent different life habits. Blue: nektobenthic (epifaunal with the ability to swim); yellow: epifaunal, brown: infaunal; red: deep infaunal; and grey: rock borer specialists.

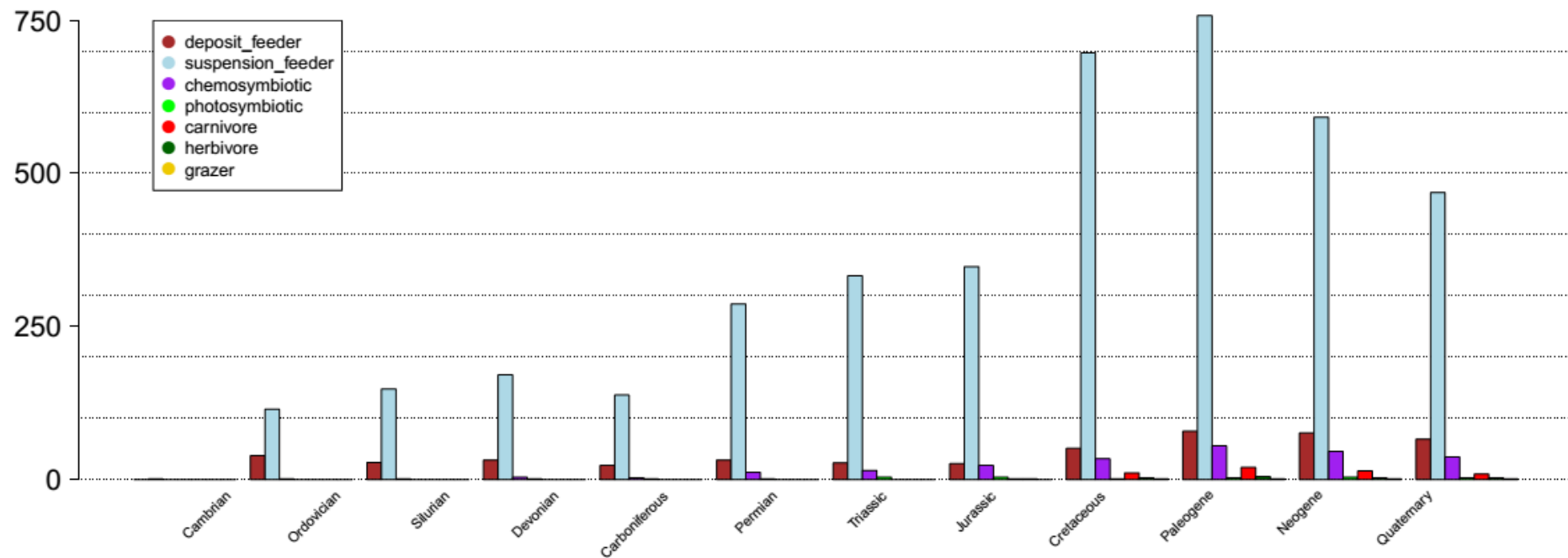


Figure.2.16 Counts of the assumed diet of bivalves across the Phanerozoic as recorded in the Paleobiology Database, split by Period. The colours represent different life habits. Maroon: deposit feeder; light blue: suspension feeder; purple chemosymbiotic; light green photosymbiotic, red: carnivore; dark green: herbivore; and yellow: grazer.

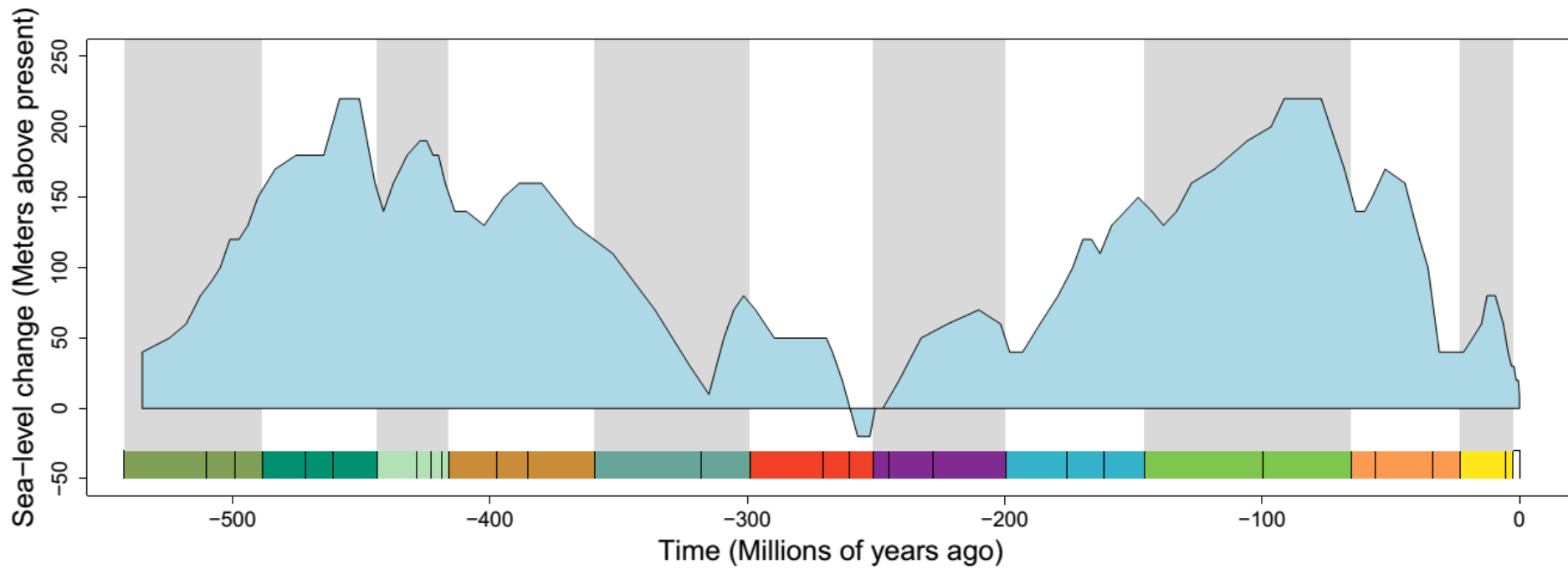


Figure.2.17 Changes in sea level across the Phanerozoic redrawn from Haq and Schutter (2002) and Snedden and Liu (2010). The values of sea level are the level above or below present day sea level (0 meters).

seen at the end of the Silurian. Through the Triassic, diversity rises, above that which is expected by the model, peaking at the End Triassic mass extinction. From this event diversity drops into the range of the expect model, but quickly rising again in the Early Jurassic, although not reaching the levels prior to the extinction event. For the rest of the period, and the Cretaceous, diversity remains within the models expected range up to the last two stages of the Cretaceous.

In these last two intervals, there a brief peak of diversity greater than expected due to the number of formations, then a dramatic drop at the boundary. This drop is the greatest of its kind (Fig.2.14). In the immediate aftermath of the extinction, diversity rises again, reaching a level of diversity above that expected by the model, and second only to modern day levels. For the rest of the Palaeogene and Neogene, the diversity falls back to within the ranges of expected diversity until present day, where diversity reaches its highest peak.

### ***2.15 Life Habit, Diet and Sea Level Change Across the Palaeozoic***

The counts of different life habits are detailed in Figure 2.15. The infaunal life habit is found to be the most common across the Phanerozoic. The epifaunal life habit was the second most common, and in much of the Palaeozoic there are statistically indistinguishable differences between the two kinds of life habit. Both life habits double in amount in the Permian but the jump was not statistically significant for either life habit (exact binomial test epifaunal:  $p$ -value= 0.160; infaunal  $p$ -value= 0.535). There is a significant departure between the number of epifaunal and infaunal life habits after the Cretaceous which remains prevalent up to present day (exact binomial test  $p$ -value= 0.0003068). The other life habits remain low in abundance for the entire Phanerozoic, gradually increasing with each Period. The counts of bivalve diet strategies through the Phanerozoic are seen in Figure.2.16. The most common diet strategy of bivalves according to the Paleobiology Database was suspension feeding followed by deposit feeding. From the Permian the numbers of chemosymbiotic genera begin to increase, eventually reaching a level statistically indistinguishable from the number of deposit feeders. The other diet strategies remain in low abundance or absent for the entire Phanerozoic.

### ***2.16 Discussion***

The Sepkoski, Treatise and Paleobiology Database sourced stratigraphic ranges agree on the general diversity count patterns across the Phanerozoic (Fig.2.1-2.4), and are congruent with published patterns in the literature (Ros & De Renzi 2005; Vinarskia *et al.* 2011; Miller & Sepkoski 1988; Dmitriev & Nevesskaja 2006; Jablonski *et al.* 2003). Diversity initially builds slowly, but the rate of increase steps up markedly after the end Permian mass extinction. This is unsurprising given that the data making up these curves

are partly modified versions of the material of Sepkoski, itself largely based on the 40 year old bivalve edition of the Treatise (Sepkoski 2002; Cox *et al.* 1971). When the data of the Sepkoski, Paleobiology Database and revised Treatise data were compared in detail, the reasons behind their congruence are less straightforward than simply sharing the same genera.

Despite all of these differences between the data, the curves remain remarkably similar to each other. The genera that are shared between the three datasets have had their ranges extended or shortened to varying degrees in the different databases; however, the overall effect on the shape of the curve seems to be minimal. The addition of new genera does not affect the overall shape of the curve as the sampling of genera is relatively even across the whole of the Phanerozoic. The congruence of shape in the face of substantial revisions of the generic ranges of bivalves suggest the trend of increasing diversity through the Phanerozoic is robust to modifications and additions to the raw data when the fossil record is read directly.

The traditional method of direct counting makes some important assumptions: (1) the fossil record provides a reliable estimate of the relative increase/decreases in diversity with time; and (2) the fossil record is reliable in the dating of the first and last appearance of genera. Both of these are big assumptions and likely false or biased in some fashion e.g. fluctuating genus abundances (more abundant genera have more chances to fossilise), fluctuating geographical ranges of genera (cosmopolitan genera have more chances for fossilisation compared to range restricted genera), fluctuating levels of fossiliferous rock (greater quantities of fossilisable sediment mean more chances to be fossilised), uneven destruction of fossiliferous rock (different rock types erode at different rates in different regions). It was for these sort of reasons that the other methods of diversity calculation were developed (Foote 2000; Alroy 2010a; Benton *et al.* 2013; Lloyd *et al.* 2012).

The shareholder quorum subsampling and sampling proxy modelling techniques were used to counter the effects of the patchy fossil rock record. The results of the shareholder quorum subsampling (Fig.2.5) are interesting as they provide a similar pattern as the traditional counts (Fig.2.1-2.4), i.e., that diversity builds slowly up to the Permian and increases thereafter; however, there are some distinct differences between the curves of these two methods. The relative magnitude of each mass extinction event differs between the traditional counts curves and each quorum level (Fig.2.5).

According to the traditional count curves of the Paleobiology Database (Fig.2.1), the End Permian, End Triassic and End Cretaceous extinctions show roughly equal drops in diversity across their boundaries, while the Treatise finds the largest drops in the End Triassic and End Cretaceous. The traditional counts of the revised Treatise (Fig.2.3-2.4) pick up an abrupt drop in the diversity across the End Triassic boundary (Rhaetian-Hettangian) whereas the Paleobiology Database (traditional and shareholder quorum)

(Fig.2.5) find the extinction to be more long ranging (Norian–Pliensbachian). The shareholder quorum results also find no evidence for a significant drop at the End Cretaceous which is evident in the traditional counts of the Paleobiology Database (Fig.2.5) and Treatise (Fig.2.1, 2.3-2.4).

These results back up current knowledge of Mesozoic diversity changes in bivalves. The time between the End Permian and End Cretaceous is considered an important interval of bivalve evolution (Ros & Echevarria 2011; Ros *et al.* 2011). Between the End Permian and End Triassic, bivalves take over from brachiopods as the dominant marine epifaunal invertebrates (Gould & Calloway 1980; Fraiser & Bottjer 2007), and explore new infaunal, siphon feeding niches absent in the Palaeozoic (Stanley 1968). After the End Jurassic, the Hippuritida (rudist bivalves) radiate. The order makes up a significant proportion of the late Mesozoic reef system, but then become extinct at the End Cretaceous.

Regardless of which set of data is looked at, traditional counts give the same relative magnitude of extinction across the Permian (50-60%). This result is surprising given the three sources only share 12 genera that become extinct at the End Permian boundary. The effects of this mass extinction on life were profound, depleting marine (Hallam & Wignall 1997) and terrestrial clades (Sahney & Benton 2008) alike. The drop in bivalve diversity at this boundary regardless of which stratigraphic source is used provides evidence that this event was an important part of bivalve evolutionary history. The extinction event is further supported by the other diversity calculation methods, which all find a drop in diversity across the boundary (apart from the range uncertainty method in Fig.2.9).

The end Triassic extinction of bivalves has been looked at in detail and is thought to have occurred quickly across the Rhaetian-Hettangian boundary with a recovery in the Sinemurian (Ros *et al.* 2011). This result is present in the revised Treatise data but the Paleobiology Database data (traditional counts and subsampling) suggests this extinction was prolonged, continuing into the middle Early Jurassic. This continuing drop in generic diversity does not contradict current estimates of fossil diversity, and instead connects the second order extinction of the Pliensbachian-Toarcian (Ruban 2006; Caruthers *et al.* 2013) to the larger mass extinction of the End Triassic.

The largest difference between the diversity estimates based on traditional counts and the subsampling procedures are found in the Late Jurassic and the End Cretaceous. The traditional counts of the Paleobiology Database (Fig.2.2) show an increase in diversity in both intervals near the end of each period, resulting in a sharp peak followed by an extinction event. The Cretaceous also has an initial burst of diversity in the early Cretaceous, followed by the second rise up to the extinction event. In the revised Treatise these diversity increases in the Late Jurassic and Early Cretaceous are not found (Fig.2.1, 2.3-2.4). The subsampled diversity of the Paleobiology Database (Fig.2.5) has the same



fluctuating pattern of the traditional counts of the Paleobiology Database, but is far less pronounced, resulting in a much flatter pattern, suggesting the extinction was minor at the generic level. As the subsampling generated curve attempts to remove the sampling bias that comes from reading the fossil record directly, I am inclined to believe the subsampling curve compared to the traditional count curve. This suggests that the bivalves were more extinction resistant at the generic level than has previously been estimated. The subsampling procedure does find the increase in diversity through the Early Cretaceous, but now suggests that this was the point of greatest diversity in the Cretaceous (Fig.2.5).

There is evidence for an extinction of bivalves during the late Jurassic in the form of the Tithonian extinction, a regional extinction of seemingly particular importance in European and Russian bivalves (Hallam 1986; Hallam & Wignall 1997). Although this event is picked up in the Paleobiology Database data (Fig.2.5) and in the published Sepkoski data (Miller & Sepkoski 1988) the few studies that have looked at this time interval disagree on when the diversity peak of the late Jurassic took place e.g. Callovian (Ruban 2006) or Kimmeridgian (Liu 1995; Crame 2002), and whether the decline proceeding it was regional (Hallam & Wignall 1997) or global (Crame 2002; Ruban 2006). The subsampling result does not remove the extinction completely from the diversity curve, but dampens its impact on generic level diversity. Although the cause of this event is still up for debate, the results of the traditional counts and subsampling suggest this decrease across the late Jurassic-early Cretaceous is a real event and not the result of sampling bias.

The Early Cretaceous rise and Middle Cretaceous crash picked up in the traditional counts and subsampling (Fig.2.5) coincides with the Aptian extinction (Hallam & Wignall 1997). This event is not evident in the revised Treatise (Fig.2.1, 2.3-2.4) but is seen in other bivalve diversity curves (Vinarskia *et al.* 2011; Miller & Sepkoski 1988). The reason behind this discrepancy between the Treatise and the Paleobiology Database (traditional counts Fig.2.1) is partly due to the inclusion of order Hippuritida (rudist bivalves) in the Paleobiology Database curve and lack thereof in the revised Treatise data. It was missing from the Treatise data as at the time of analysis and writing up of this thesis the stratigraphic ranges of the order were undergoing revision and not available for analysis. This missing order from the revised Treatise is not ideal but does not pose a significant problem for two reasons: (1) the order itself only occurred in the Cretaceous and so their contribution or lack thereof only effects the diversity pattern during that period ; (2) their removal from the Paleobiology Database data does not change the pattern of diversity through the Cretaceous, suggesting the changes in diversity though this interval is not due to the presence of rudists alone.

The high origination rates of bivalves through the early Cretaceous can explain the increase in diversity up to the Aptian (Nurnberg & Aberhan 2013), but during the extinction and the time interval thereafter, low origination and extinction rates cannot explain the

increase across the rest of the Cretaceous. The Aptian-Turonian time interval has been well studied, and it is generally agreed that it was a time of global warmth, high sea levels and high tropical water surface temperatures (Norris *et al.* 2002; Hay & Floegel 2012), with several episodes of ocean anoxia (Leckie *et al.* 2002), but was also visited by at least one cold snap (Mutterlose *et al.* 2009).

The most interesting result of the subsampling curve (Fig.2.5) compared to the traditional diversity curves of the three databases (Fig.2.1) is the reduced magnitude of the Cretaceous mass extinction. The bivalves most affected by this extinction event were the Hippuritida as the clade became extinct at the end of the Cretaceous. Although they are absent from the Treatise data, the extinction event is still highly visible (Fig.2.3-2.4), and their removal from the Paleobiology Database does not change the result of Figure 2.5. The subsampling curve result suggests that impact of the extinction on bivalves was small, and mostly an artefact of sampling biases. If this is indeed the case, then one could expect the sampling proxy modelling results to find the diversity at this time also to be a sampling artefact; however, as Figures 2.13-2.14 show this does not seem to be the case. The results of Figures.2.11 & 2.13 (using collections as a sampling proxy) show that during the event (i.e. Maastrichtian-Danian) the observed diversity rapidly shifted from significantly lower diversity than expected to much greater.

The results of Figures.2.12 & 2.14 (using formations as a sampling proxy) show the same pattern, but this switch begins slightly earlier (the Campanian), suggesting that through the event the diversity in the Maastrichtian was purely the result of the number of formations. The subsampling results (Fig.2.5) and the modelling result using formations (Fig.2.14) suggest that the diversity change at the extinction event was mostly the result of sampling, while in the stage after the event (Danian) was found to contain greater diversity than expected given the number of formations or collections. The results of the subsampling and modelling approaches are compatible. The subsampling result shows that the change in diversity across the boundary was much smaller than in traditional diversity curves (Miller & Sepkoski 1988; Jablonski *et al.* 2003; Ros & De Renzi 2005; Kidwell 2005; Dmitriev & Nevesskaja 2006; Nevesskaya 2008; Vinarskia *et al.* 2011).

This is backed up by the modelling approaches (Fig.2.13-2.14) which suggest that the diversity of the Maastrichtian was low with respect to the number of collections, and/or entirely the result of the number of formations available. The modelling approaches also suggest that the diversity after the event was much greater than expected given the sampling proxies. For the Cretaceous extinction the results here suggest that the diversity drop across the End Cretaceous boundary was much lower than previous diversity curves have found. The literature which has looked at the extinction of bivalves across this boundary in detail suggest that the bivalve extinctions at this time were phylogenetically conserved (Roy, Hunt & Jablonski 2009) and that the intensity of the mass extinction fell away from the tropics (Vilhena *et al.* 2013). If phylogenetic conservatism was the reason

for the extinction of bivalve genera at this boundary, one would expect to find these genera belonging to the same family (or higher group). This does not seem to be the case here, as those taxa becoming extinct are from different families. Distance from the tropics does not seem to explain this result either, as the taxa involved were from a range of tropical or extra-topical habitats (Kiessling *et al.* 2013).

The traditional counts and shareholder quorum subsampling curves for the Palaeozoic show the same overall trajectory i.e., compared to the whole diversity trajectory, the Palaeozoic had low diversity and was relatively stable. The Ordovician mass extinction is seen in the traditional counts (for bivalves and the three subclades) but does not get picked up by most of the SQS results (Fig.2.5, 2.7-2.8). The exception is in the SQS of Fig.2.6 showing the results of the Protobranchia. The sampling proxy results also show a significant departure from the diversity expected using either numbers of collections or formations (Fig.2.13-2.14). The departure was negative; suggesting the number of genera in the interval was much lower than expected. In all cases the drop in diversity begins before the End Ordovician mass extinction boundary, starting in the Middle Ordovician.

A source of known bivalve extinction during the late Ordovician were the Ashgill (Middle Katian-Hirnantian) glaciations, which brought about a reduction in sea level (Finnegan *et al.* 2011). The extinctions themselves were mostly of epifaunal and semi-infaunal clades in low latitudes (Cope 2002), and mostly of taxa within the Protobranchia and Heteroconchia (Cope 2004). The SQS results of the analysis agree with the literature for the relatively large drop in the Protobranchia (Fig.2.6) through the boundary but the decrease in Heteroconchia was more subtle (Fig.2.7). Through the Ordovician, the change in diversity according to Cope (2004) was roughly equal to that of the Protobranchia suggesting the impact of the mass extinction on heteroconchs was less than previously published. The results of the traditional counts, SQS and sampling proxies modelling agree that the Ordovician mass extinction had some effect on bivalves, but this seems to have been mostly due to a drop in Protobranchia (Fig.2.6). The cause of the extinction cannot be directly tested with the data herein, but the literature suggests that the movement of the continents, i.e., the shift of Gondwana toward the poles and drop in sea level (Haq & Schutter 2002; Snedden & Liu 2010) (Fig.2.17) continued with the growing of the Gondwanan ice sheet, which may have contributed at least in part to the extinction (Cope 2002; Cope 2004; Finnegan *et al.* 2011). The timing of the continents shifting and formation have a broad enough range (Middle-Early Silurian) and so fit with the diversity trajectories found here (Fig.2.1-12)

The Silurian rise in diversity was seen in the traditional and SQS results of the bivalves and constituent subclades. In Fig.2.13 there was a small positive deviation outside of the area assumed to be the result of the sampling bias proxy (number of collections), suggesting that the diversity peak in the Silurian was greater than expected.

The same pattern is not found in Fig.2.14, which does not find the diversity to be outside the range predicted by the number of formations. The rise in Silurian bivalve diversity could be the result of the higher origination rate over extinction rate suggested to have occurred in the Silurian marine biota (Bambach *et al.* 2004).

During the Silurian sea level increased (Haq & Schutter 2002; Snedden & Liu 2010) (Fig.2.17) giving way to increases Palaeozoic reefs (Munnecke *et al.* 2010); however, some evidence suggests that Silurian reefs were not sites of high bivalve diversity (Watkins 1997). The diversity rise in bivalves during this Silurian is interesting as some Phanerozoic authors have shown a similar increase (Alroy 2010c; Vinarskia *et al.* 2011; Miller & Sepkoski 1988) while others have not (Nevesskaya 2008). No studies have emerged that have looked at the Silurian record of bivalve diversity in detail to pick apart the reasons for the increase, and so this would make a good area for future work.

The Devonian mass extinction event is not picked up as a singular drop in diversity in any of the diversity curves (Fig.2.1-13). Instead the diversity of bivalves seems to drop steadily in all subclades up to the end of the Devonian boundary. This falls in line with the results of previous diversity studies that relied on traditional counts (Miller & Sepkoski 1988; Vinarskia *et al.* 2011; Nevesskaya 2008). The results of Fig.2.13 show a large negative deviation from the diversity expected of the number of collections, but only a small negative deviation from the number of formations results (Fig.2.14). Therefore some of this drop in diversity could be partly an effect of the number of formations available, but not for the whole of the Devonian. The latest Phanerozoic diversity curve of Alroy *et al.* (2010c) showed a similar drop in diversity in other benthic marine clades (i.e. Trilobites, Brachiopods and Anthozoa).

Sea level drops at the beginning of the Devonian, rising in the Middle Devonian before dropping again (Fig.2.17) and continuing to fall for the rest of the Devonian and Carboniferous (Haq & Schutter 2002; Snedden & Liu 2010). At the same time the continents of Baltica and Laurentia collided together (beginning in the late Silurian) forming Euramecia, closing the Iapetus Ocean. This event could explain the gradual decline in diversity across the Devonian. Looking at the Devonian bivalve diversity record in more detail in future projects could help identify key intervals or events that resulted in the patterns seen in this section, and which may help to better understand the patterns of diversity decline in other marine organisms such as the brachiopods or stromatoporoid corals.

The last part of this analysis was to look at the changes in diversity of the three subclades: Protobranchia, Heteroconchia and Pteriomorphia. The three subclades (Fig.2.6-2.8) have very different diversity profiles through the Phanerozoic. The Protobranchia shows a very erratic trajectory compared to the more densely sampled Heteroconchia and Pteriomorphia. This will be partly due to the low levels of Protobranchia diversity through the Phanerozoic; however, the SQS results found the

same pattern of high diversity in the Early-Middle Palaeozoic (Ordovician-Devonian), stable trajectory across the rest of the Palaeozoic and Mesozoic and increase again in the Cenozoic. The Heteroconchia and Pteriomorphia (Fig.2.8) appear in the Ordovician but are known to diversify after the End Permian extinction (Cope 2000); however, the new data (traditional counts, Paleobiology Database) suggest they started to rise in diversity in the Permian. The origin of the deep infaunal siphon feeders appear (Stanley 1968) during this time which could provide some explanation; however, since this original paper, no revision of the study has been carried out to confirm if this result is still sound. The counts of life habit and diet (Fig.2.15 and 2.16) show that by the end of the Ordovician, that epifaunal to deep infaunal habitat use had evolved, which fits with previously published work (Sanchez 2008), but that the majority of taxa feed upon the suspension. Currently the Paleobiology Database does not contain information on the type of siphon used by the bivalves therefore work still needs to be done before the database is useful for testing the claims of Stanley (1968).

## **2.17 Conclusions & Future Prospects**

The aim of this Chapter was to revise the bivalve (and subclades) diversity curve using the revised ranges of the Treatise and Paleobiology Database and to construct new curves using the most up-to-date techniques available in the paleontological literature. Using these new techniques I have shown that the bivalve diversity trajectory through the Phanerozoic is relatively robust to new data and sampling biases. The data also provided evidence that one extinction event important in other clades i.e. the End Cretaceous mass extinction, did not affect the bivalve generic diversity curve significantly. The end Triassic extinction was an important event for bivalves, but extended beyond the End Triassic into the Early Jurassic (Toarcian). The Permian extinction left its mark on bivalve diversity like many other clades but its impact was felt more in Heteroconchia and Pteriomorphia than Protobranchia (which increased across the end Permian). The Ordovician extinction affected the Protobranchia more than the Heteroconchia and Pteriomorphia while all three clades gradually fell across the Devonian.

From this data new avenues of enquiry open up. The reasons behind the changes in diversity through the Palaeozoic need to be looked at in more detail. There is evidence that the diversity trajectories for bivalves are a mosaic of changes in latitudinal diversity (Cope 2002), but this study did not track latitude. The factors behind individual events, such as the increase in diversity across the Silurian, and drop gradual drop in the Devonian despite a sea level rise through the Middle Devonian should also be explored. The analyses of the chapter looked at the changes in diversity of subclades at the generic level, but as seen in the studies of Ros and De Renzi (2005) and Nevesskaya (2008), the techniques applied in the Chapter could also be applied to other taxonomic levels. This study chose the generic level to use the three sources of stratigraphic data (Sepkoski,

Treatise and the Paleobiology Database), but with the new bivalve classification (Carter *et al.* 2011) future work could test what effect taxonomic level has on bivalve Phanerozoic diversity trends using the same diversity techniques as used in this chapter.

The comparison of the Treatise and Paleobiology Database data showed that despite differences in the ranges, the same overall pattern of diversity increase is produced by traditional counts and SQS. One problem with the stratigraphic data in its current form is the inconsistencies for shared genera. The next step for future analysis using these data should be to go through in detail comparing and contrasting the ranges of each genus in common to pick out the major inconsistencies. From this position one can build the most up-to-date bivalve stratigraphic dataset. The dataset chose to look at the diversity based on traditional counts, subsampling and modelling techniques; future work should take this data and use it to build detailed origination/extinction curves. These can then help to assess whether peaks and troughs in the diversity curve are due to increased/decreased origination and extinction rates.

Knowing the rate of origination and extinction in bivalves across the Phanerozoic, when combined with stratigraphic and latitude information, will help to identify areas that generate diversity that supply other areas (macroevolutionary sources) and help to explain how they are formed/destroyed (Jablonski *et al.* 2006). One would expect that given that the tropics are considered the macroevolutionary source of diversity for bivalves in higher latitudes for the last 11 million years (Jablonski *et al.* 2006) that this would be true further back in time. One would also expect that for bivalves the latitudinal centre of thier macroevolutionary source would have migrated from the southern higher latitudes of Gondwana where bivalves are proposed to have originated (Sanchez 2008) towards the tropics latitudes of the equator as bivalves explore more of the globe with the movement of the continents and opening/closing of oceans.

## ***Chapter 3***

### ***The Palaeozoic Disparity of Bivalves***



### 3.1 Introduction

While the previous chapter inferred diversity curves for bivalves throughout the Phanerozoic, the present chapter focuses on the Palaeozoic to explore the disparity of bivalves during the first half of their evolutionary history. The published literature provides a general account of the changes in bivalve diversity across the Palaeozoic (Ros & De Renzi 2005; Vinarskia *et al.* 2011; Miller & Sepkoski 1988; Dmitriev & Nevesskaja 2006; Jablonski *et al.* 2003). The Palaeozoic is a time when nearly all Phyla made their first appearance in the fossil record (Valentine *et al.* 1999) and subsequently radiated (Cope 2004). Through the Cambrian five basal bivalve genera appear (Elicki & Gursu 2009), belonging to the Euprotobranchia (Carter *et al.* 2011), a grade of bivalves that do not extend beyond the Cambrian period. Bivalves belonging to the Eubivalvia (i.e. all other bivalves) initially diversified in the Early Ordovician (Cope 2004), beginning to explore epifaunal/shallow infaunal non-siphonate guilds in marine and estuarine habitats (Sanchez 2008). During this radiation there were four major Ordovician landmasses: Gondwana, Laurentia, Baltica and Siberia, all of which were located mostly within the southern hemisphere. Bivalves are initially found in Gondwanan derived sediments but by the Middle Ordovician they had found their way to Laurentia and Baltica, presumably a consequence of evolving planktonic larvae (Sanchez 2008; Polechova 2013).

During the Silurian and Devonian, the three continents Laurentia, Baltica and Siberia collided to form Euramerica which was spread across the tropical latitudes of the southern and northern hemisphere. At this point Gondwana still occupied a significant portion of the southern hemispheres extra-tropical latitudes. By the Carboniferous the continents were beginning to collide into one another, forming the super continent Pangaea by the end of the Permian. While the continents shifted, the bivalves continued to grow in diversity and in geographic range (Ros & De Renzi 2005; Vinarskia *et al.* 2011; Miller & Sepkoski 1988; Dmitriev & Nevesskaja 2006; Jablonski *et al.* 2003), making their way north (Naimark & Markov 2011). Up to the end of the Palaeozoic they played second fiddle to the brachiopods in diversity and abundance (Fraiser & Bottjer 2007), but were able to dominate some marine communities (Miller 1989; Kriz 2004). Despite their similarity, the diversity trajectories of these two clades through the Phanerozoic provides little evidence for the competitive replacement of bivalves over brachiopods (Gould & Calloway 1980).

The understanding of bivalve evolutionary history has been almost entirely reliant on the taxonomic changes in diversity through time. The results of Chapter 2 suggest the increase in bivalve diversity through time is a robust trend, but the strength of individual events such as the late Devonian or End Cretaceous mass extinction can vary depending on the method of diversity calculation. Morphological disparity analyses provide a complementary source of data for making sense of patterns in diversity (Chapter 1) but for bivalves there



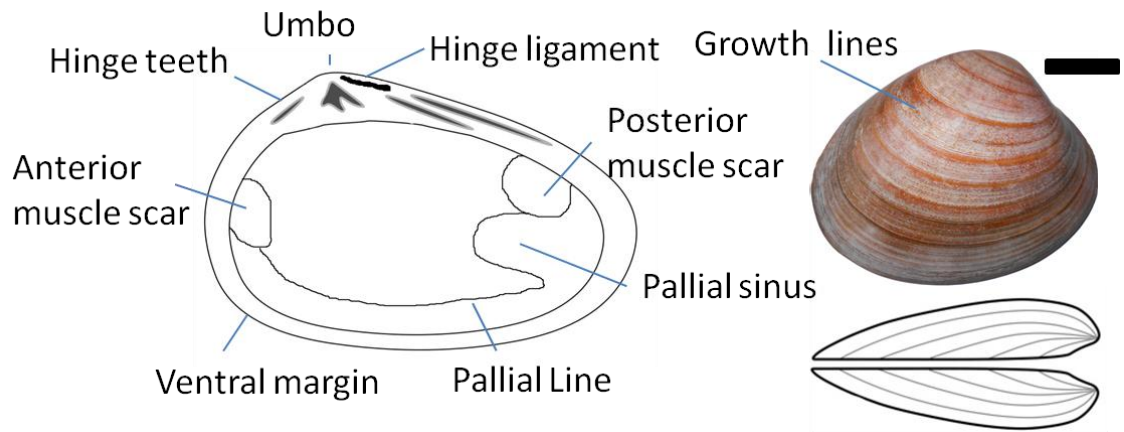


Figure.3.0. The internal structures of the right valve, the external anatomy and side view of the two valves together. Black scale bar represents 1 cm.

exist only small, isolated studies for the Mesozoic-Cenozoic and no published record of bivalve disparity through the Palaeozoic. This chapter fills this significant gap by quantifying the morphological disparity of bivalves in the Palaeozoic.

### 3.2 The Bivalve Body Plan

The evolutionary success of bivalves comes down to their hardy body plan (Fig.3.0-3.2). With it bivalves have adapted to numerous marine and freshwater habitats and occupy various ecological niches (Fig.3.2) (Stanley 1968; Sanchez 2008). The most prominent feature that distinguishes them from other molluscs is the pair of hinged calcareous shells (valves) that encapsulate the animal. The valves themselves come in various shapes, from delicate elongated razor shells (e.g. *Ensis arcuatus*) and fan like scallops (e.g. *Chlamys hasta*) to the robust forms of cockles (e.g. *Acanthocardia*) and oysters (e.g. *Crassostrea gigas*). The majority of species are separated on internal characters such as the size/shape/position of the pallial sinus (attachment point of the siphons) and muscle scars (Stanley & Yang 1987) (Fig.3.0).

The hinge itself is a mix of diagnostic characters, including the fibrous/lammellar proteinaceous ligament (Zengqiong & Gangsheng 2011) and interlocking hinge teeth (Ubukata 2003). The hinge teeth come in a variety of shapes, from simple rows of uniformly shaped teeth (taxodont) to forms of asymmetrically aligned, heterogenic shaped teeth (heterodont, desmodont and schizodont) (Fig.3.1A). The hinge ligament, the structure that opens the valves on the relaxation of the internal muscles also takes on many forms. The

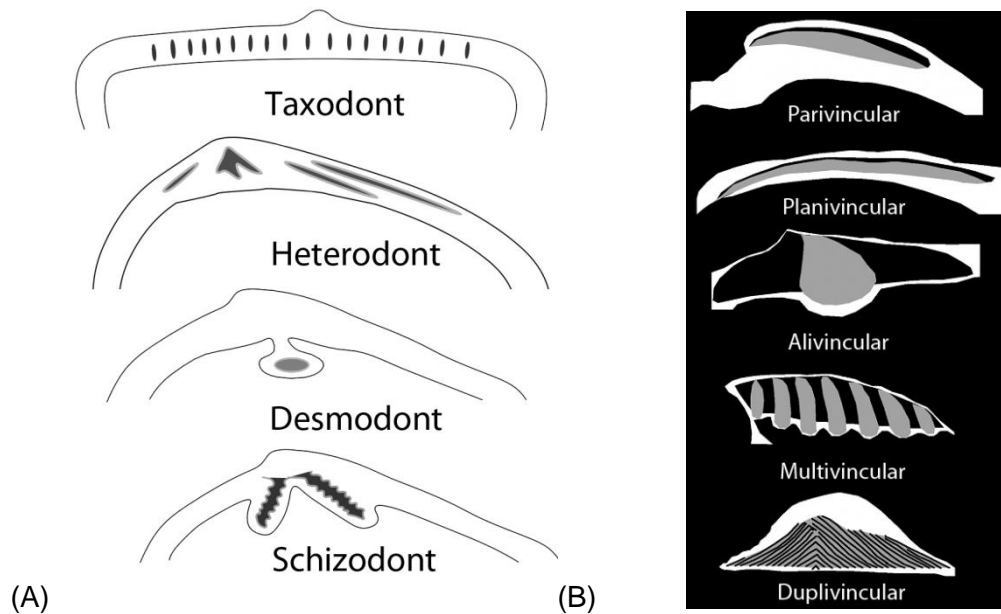


Figure.3.1. Range of hinge dentition and ligament shape/ structure. (A) Hinge teeth structure. The black and grey areas are the shape and location of the teeth on the right valve. (B) Ligament structures and shape within the hinge, black and grey marks represent the lamellar and fibrous layers of the ligament; modified from Ubukata (2003).

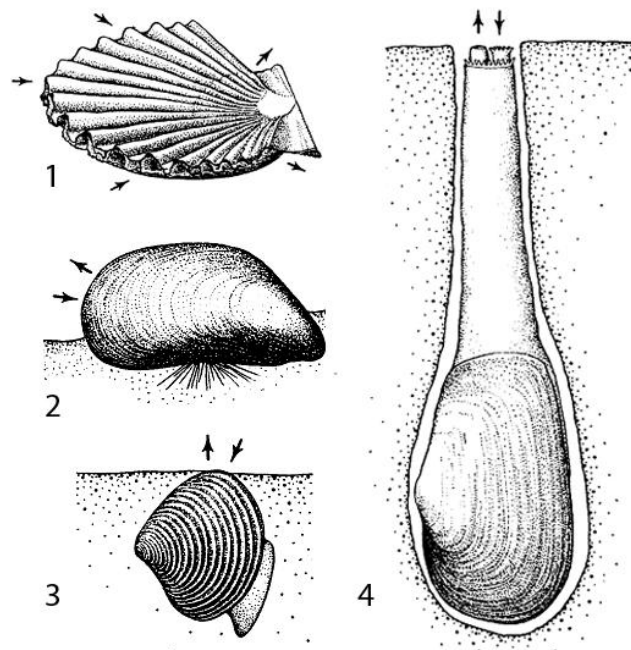


Figure.3.2. Range of life habits of marine bivalves. (1) Epifaunal and with ability to swim (*Pecten*); (2) Epifaunal and stationary (*Mytilus*); (3) Shallow infaunal (*Astarte*); (4) Deep infaunal (*Pholas*). Arrows show direction of water flow drawn in from siphons Diagram modified from Stanley (1968).

simplest are the elongated, singular layers of fibrous and lamellate ligament supported (parivincular) or unsupported (planivincular) by a platform structure secreted as part of the valve (Ubukata 2003). Alivincular hinge ligaments are slightly more complex, the lamellar element of the ligament is split in two, with the fibrous layer found in the middle, taking on a triangle shape (Fig.3.1B). Multivincular ligaments are made of repeating layers of fibrous and lamellar elements of the ligament (Fig.3.1B). Duplivincular ligaments are also made

up of layers of repeating fibrous and lamellar elements; but these are externally secreted below the umbo (Fig.3.0) giving a chevron like appearance (Fig.3.1B) (Ubukata 2003).

Bivalves share the muscular foot of other molluscan clades, but can be distinguished from those clades by the reduction or modification of the foot into a digging tool. They are unique among the molluscs in using their siphons as a means of absorbing oxygen and other essential molecules from the water, and in more derived forms, as a means for feeding and respiration (Autobranchia) (Carter *et al.* 2000). The size and length of the siphon is indicative of the habit and mode of life. Longer siphons allow animals to bury themselves deeper in the sediment (Fig.3.2). It is these characters, as well as a raft of genomic data that confirm the monophyly of the group (Giribet & Wheeler 2002; Kocot *et al.* 2011; Plazzi *et al.* 2011).

With this simple yet diverse body plan, bivalves display a rich ecological variability, which has been the case since the early Ordovician (Sanchez 2008). Modern bivalves inhabit marine and freshwater habitats across the globe, from high energy environments such as tropical reefs and hydrothermal vents (Duperron *et al.* 2013) to the low energy environments of polar oceans (Peck *et al.* 2007). Their occupation and ecological role within these environments are highly variable (Etter *et al.* 2005). Some species (e.g. the common mussel *Mytilus* or common clam *Pecten*) have adapted to the sedimentary surface (epifaunal), while others (e.g. *Astarte* or *Pholas*) burrow into the sediment (infaunal) to varying degrees (Fig.3.2). Some species have gone one step further, boring into solid rock or coral (e.g. *Botula fusca*) (Valentich-Scott & Dinesen 2004) or have the ability to actively swim through the water column as an adult (e.g. *Pecten* or *Chlamys*) (Donovan *et al.* 2004) (Fig.3.2). They also occupy a variety of niches, from active/inactive suspension feeders or detritivores (Navarro *et al.* 2008) to carnivores (Ockelmann & Dinesen 2011).

### **3.3 Disparity in Bivalves**

Considering their lengthy fossil record and reasonable level of morphological variation, bivalves would seem an excellent study clade for morphological disparity analyses. To date these analyses have been limited to just a handful of recent case studies on relatively small post-Palaeozoic subclades (Appendix.1.Table.1.0). One investigation addressed the effect of the Pliocene-Pleistocene extinction on the disparity trajectories of veneroid bivalves (edible clams and cockles), part of the Veneroida (Kolbe *et al.* 2011) which belongs to the infraclass Heteroconchia (Carter *et al.* 2011). The authors quantified morphology using landmarks, and calculated disparity as the mean squared Procrustes distance between specimens and their mean form (Zelditch *et al.* 2004). By looking at the difference in disparity between pairs of taxa across the extinction boundary (those who survived vs. those who became extinct) the study found those

veneroid taxa surviving the extinction were more disparate than those taxa that went extinct.

A similar result was found in the disparity of teleost fish that survived the End Cretaceous mass extinction (Maastrichtian - Danian boundary) (Friedman 2010). These two studies suggest that extinction events could increase the disparity of clades fortunate enough to survive them. Bivalves have survived all mass extinction events throughout their 542 million year history. If mass extinction can increase levels of disparity by filtering out those clades that are less disparate, one may see an increase in disparity in the aftermath of each extinction event. Bivalves make an excellent dataset with which to test this hypothesis.

Another case study looked at the morphospace occupation, built using discrete characters, of Miocene corbulid bivalves (basket clams) (Anderson *et al.* 2010), belonging to the family Corbulidae which rests in the infraclass Heteroconchia (Carter *et al.* 2011). The study produced a strict consensus maximum parsimony tree with the same characters used to build the morphospace, and plotted the habitat information of these bivalves along the tree and through the space. The authors used this information to classify the lacustrine Miocene Pebas formation as a cradle of diversity (Anderson *et al.* 2010) as it was at this locality that several freshwater species of corbulid became adapted to lacustrine environments.

The “cradle” concept, i.e., an area where diversity is generated, sourcing the surrounding areas, and other associated concepts (i.e. museums, casinos, sinks, or sources) for diversity are well over four decades old (Stebbins 1974) and have been applied to other bivalve containing formations (Sanchez 2008). Prior to this paper, the concept of the cradle of diversity was based on taxonomic data alone, using origination and extinction rates across latitude (Jablonski *et al.* 2006; Roy & Goldberg 2007; Moreau & Bell 2013; Arita & Vazquez-Dominguez 2008).

The study of Grey *et al.* (2010) chose to explore the changes in disparity of buchiid bivalves across the Late Jurassic – Early Cretaceous boundary. The study used selected measurements of the shell of the genus *Buchia*, a member of the Buchidae which belongs to the infraclass Pteriomorpha (Carter *et al.* 2011). This particular boundary was chosen because it has been identified as a period of regional, infaunal bivalve extinction (Chunlian 2000). The genus *Buchia* itself was chosen as it has been a useful biostratigraphic marker due to its abundance and geographic range through the interval (Grey *et al.* 2010). This paper was innovative in plotting the disparity of bivalves across latitude, which has only been touched upon rarely over the last two decades (Grey *et al.* 2010; Roy *et al.* 2001; Neige 2003); however, this case study only looked at one time interval in a specific genus, and therefore the disparity across latitude in bivalves it is still totally unknown for the rest of the Phanerozoic.

The case study of Schneider *et al.* (Schneider *et al.* 2010) focused efforts on cataloguing the changes in shape of three genera: *Arcomytilus* (family Mytilidae) and *Isognomon* (family Pteriidae) which belong to the infraclass Pteriomorpha, and *Eomiodon* (family Neomiodontidae) which belongs to the infraclass Heteroconchia, over the Upper Jurassic. Measurements of the body form were made, but outlines were used in the construction of theoretical morphospaces of each separate genus. Taxa were then grouped by lithostratigraphy using convex hulls. By combining these data the authors found that in the Kimmeridgian (Upper Jurassic) all three genera increased in size but did not change dramatically in outline. The genus *Arcomytilus* and *Isognomon* valves became slightly more triangular in this interval, while in *Eomiodon* the change was toward more elongate shapes.

Unlike the case study of Grey *et al.* (2010) which included the entire range of the genus, Schneider *et al.* (2010) only looked at a very specific period of time (Late Oxfordian – Late Tithonian) for genera with longer stratigraphic ranges. The study (Schneider *et al.* 2010) looked at each genus individually to distinguish morphological variability among populations, finding only very minor changes in outline. By studying the genera within the same morphospace, the authors may have distinguished those environmental conditions shared by the genera causing the minor changes in shape, and find those potential shapes not explored by all three taxa.

Although these case studies have contributed to a fuller understanding of bivalve evolution, they have also left significant gaps in current knowledge. The simplest and most fundamental questions concern the manner in which bivalve disparity changed across longer intervals of time and what does this tell us about bivalve evolutionary history. The reason why these questions have not been addressed in bivalves until this thesis is likely due to the amount and kind of morphological data available to conduct these kinds of analyses. The largest bivalve dataset made with disparity analysis in mind is that of Grey *et al.* (2010), which has 1855 specimens, but whose scope was restricted to a single genus across a small period of time. The largest disparity dataset to date (1195 crinoid species) has taken at least half a decade to reach its current level (Foote 1994a; Foote 1999) and is still undergoing analysis and revision (Deline & Ausich 2011). Combining the data of the bivalve disparity analyses is inappropriate for the same reasons as those given in Chapter 1 for not using the current disparity literature in a meta-analytical fashion, i.e., a mixture of data types, disparity measures, taxonomic and stratigraphic scope (Hughes *et al.* 2013; Anderson *et al.* 2010; Grey *et al.* 2010; Schneider *et al.* 2010; Kolbe *et al.* 2011).

In addressing whether disparity was an evolutionary rule in Chapter 1 (Hughes *et al.* 2013), a suitable bivalve dataset (Carter *et al.* 2000; Carter *et al.* 2006) was acquired that can now give a partial answer to these questions. It was suitable as it used discrete characters that capture the morphological variation of a greater range of organisms than other data types; i.e. measurements, outlines and landmarks (Wills 2001); and its size of

242 genera makes the data one of the widest selecting bivalve focused morphological datasets in the literature (Carter *et al.* 2000; Carter *et al.* 2006). The analysis of Chapter 1 concluded that bivalves show a significantly top-heavy pattern of disparity across the Phanerozoic (Hughes *et al.* 2013). This is an interesting result as bivalves first appear in the Cambrian and have utilised the same basic body throughout their evolutionary history. The main conclusion of Chapter 1 was that early-high disparity is the most common trend, but bivalves do not show this. This result is more even more interesting when it is considered that the original objective behind the paper which spawned the bivalve data was to address the basal relationships of the bivalves (Carter *et al.* 2000; Carter *et al.* 2006).

Given this bias in the bivalve data toward the Palaeozoic, this dataset (Carter *et al.* 2000; Carter *et al.* 2006) was worth looking at in greater detail than the broad brush treatment it was given in Chapter 1. Not only did this provide a greater understanding of the evolution of disparity of Palaeozoic bivalves, for which there is no published information, the data also contributes to the body of literature on the disparity of Palaeozoic invertebrate clades, i.e., the Echinodermata (Foote 1999; Deline & Ausich 2011; Lefebvre *et al.* 2006; Gavrillets 1999), Arthropoda (Foote 1993a; Webster 2007; Briggs *et al.* 1992; Wills *et al.* 1994; Wills 1998b), Mollusca (Wagner 1997; Wagner 1995; Saunders & Work 1996), and Priapulida (Wills 1998a; Wills *et al.* 2012). There are a number of questions that naturally follow on once the main disparity pattern of bivalves has been established and these are discussed.

A handful of studies have shown that disparity patterns of clades can be a combination of various subclade patterns (Foote 1993b; Brusatte *et al.* 2012; Lupia 1999; Foote 1999; Anderson *et al.* 2011). How much do the individual subclades of the Bivalvia contribute to the disparity pattern? The published disparity analyses (Anderson *et al.* 2010; Grey *et al.* 2010; Schneider *et al.* 2010; Kolbe *et al.* 2011) have focused their efforts on individual genera within the Heteroconchia and Pteriomorphia; even if it was assumed that these genera were typical of their represented families their combined taxonomic range covers <10% of the known bivalve families, and does not cover any families of the Protobranchia, the sister subclass to the Autobranchia (Heteroconchia plus Pteriomorphia) (Carter *et al.* 2011).

The way in which bivalves evolve through the morphospace is unknown for the Palaeozoic, and has only been lightly touched upon in the Mesozoic-Cenozoic (Anderson *et al.* 2010; Grey *et al.* 2010; Schneider *et al.* 2010). The way in which the taxa cluster in morphospace can give some context to the changes in disparity by showing whether taxa are occupying similar areas of the space or going in different directions i.e. converging on the same morphological phenotypes or evolving away from each other. Even if the taxa are evolving in different directions, their trajectories could be random, static, or because of adaptation (gradual or rapid).

These explanations can all be scrutinised using evolutionary models (Hunt *et al.* 2008; Geary *et al.* 2010). These models have yet to be put to disparity patterns directly but can help to distinguish whether disparity trends are the result of: an unbiased random walk, generally associated with genetic drift; a general random walk (unbiased walk where the mean step change is not zero), associated with directional change; evolutionary stasis (uncorrelated fluctuations around a static mean), associated with no change; and the Ornstein-Uhlenbeck process (random walk with a an attracting optimum), associated with adaptive radiation (Hunt 2006; Geary *et al.* 2010; Butler & King 2004).

An important follow-on question that should be addressed is what effect does mass extinction have on the long term evolution of disparity within bivalves? As stated above, published disparity papers (Friedman 2010; Kolbe *et al.* 2011) have shown that extinction survivors can be more disparate than their extinct counterparts. As there have only been two studies looking specifically at the change in disparity across mass extinction boundaries (Friedman 2010; Kolbe *et al.* 2011), there is still much uncertainty whether this is the general rule across the Metazoa. If true it may mean reevaluating how we perceive mass extinctions which have commonly been seen as catastrophic events with little benefit to the continuation of biodiversity. Bivalves have survived all mass extinctions to date. The Palaeozoic bivalve data of this section provides an excellent opportunity to explore how mass extinctions affect the long term evolution of disparity profiles as the Era contains two of the big five mass extinctions (End Ordovician and Late Devonian). By looking at the change in disparity across these two boundaries one can begin to build up a picture of how different mass extinctions have affected the disparity profiles of a single clade.

### **3.4 Materials & Methods**

#### **3.5 Data Collation**

The morphological data used in all of the disparity analyses presented herein was taken from the discrete morphological character matrix of Carter *et al.* (2000; 2006) (supplied by the author). The matrix contained 254 discrete characters. Of these, 246 characters described shell structure and shape, seven characters described different aspects of the mantle, and one character reported an outdated and approximate first appearance for each genus in the fossil record. The stratigraphic range character was removed as it does not provide any morphological information. There were a total of 8 possible Palaeozoic originating genera that extend to the present according to the Paleobiology Database or the Treatise. As the numbers of soft body characters were few in number and only arose in extant genera, their effect on the construction of the Palaeozoic morphospace was negligible and so were left in.

The original matrix contains 350 bivalve specimens and 4 outgroup taxa: *Watsonella crossbyi* and *Anabarella plana* (Monoplacophora), *Ribeiria junior*

(Rostroconchia), and *Pseudomyona queenslandica* (Helcionelloida) (Parkhaev 2006). Depending on the stratigraphic source, these 346 bivalve specimens in turn consisted of 198 or 185 Palaeozoic originating genera (Mixed Range and Treatise stratigraphy respectively), 30 or 42 Mesozoic and 14 Cenozoic originating genera. Specimens of the original matrix were selected to sample from as many of the Palaeozoic (and to a lesser extent Mesozoic) families as possible to build a matrix representing the major bivalve groups of the time (Carter *et al.* 2000; Carter *et al.* 2006). Specimens that came from small, poorly known families built upon damaged/poor quality material were omitted to make the analysis more practicable. This should not affect the calculation of total Palaeozoic bivalve disparity as those genera that remain form the bulk of Palaeozoic bivalve diversity. The Mesozoic and Cenozoic genera were part of the original data to benefit tree construction, i.e., bridge the gap between the Palaeozoic originating taxa and those which originated later. As the disparity analyses focused on the Palaeozoic, all genera that did not originate in that Palaeozoic interval were removed.

Even when only looking at Palaeozoic, at the species level some genera were oversampled relative to others due to practical constraints i.e. inability to find multiple species of specific taxa. Variance-based indices of disparity are sensitive to this type of sampling bias because it assumes that the selection of taxa was not the result of artificially uneven sampling (Wills 2001; Ciampaglio *et al.* 2001). If some groups are artificially over or undersampled compared to others, the effect on disparity will be a reduction in apparent variance of the more intensively sampled groups. The centroid of the empirical morphospace would also be pushed towards the more intensively sampled groups if they occupy the same area of morphospace. To reduce the effect of inhomogeneity of sampling, specimens were amalgamated to the rank of genus. This ranking was chosen over family or order as a compromise between: (1) retaining the finest morphological resolution between individual taxa, (2) reducing the need for a-priori assumptions on the inclusion of taxa in higher taxonomic ranks undergoing revision (Carter *et al.* 2011), and (3) coinciding with the taxonomic resolution of data mined from the Treatise of Invertebrate Palaeontology and Paleobiology Database.

In all analyses, outgroup taxa were removed. In disparity analyses outgroup taxa are inappropriate as they are usually undersampled with respect to the sampled ingroup, i.e., undersampling the taxonomic and morphological variation of lineages close to the root of the total clade (ingroup + outgroup). Morphospaces containing outgroup+ingroup taxa may also not make sense if the outgroup is morphologically and/or taxonomically far removed from the ingroup. The taxonomic scope of the original matrix captures a range of morphologies explored by three major subclades of bivalve: the Subclass Protobranchia (equivalently known as Palaeotaxodonta and used interchangeably to refer to the same clade), and Infraclasses Heteroconchia and Pteriomorphia (Subclass Autobranchia) (Carter *et al.* 2006; Carter *et al.* 2011). The three clades combined make up the known



Clade	Temporal Range	Condense Level	Specimen Count	Total Character Count	Characters Pruned	Remaining Characters	Net loss
Bivalvia	Phanerozoic	Specimens	350	253	-	-	-
Bivalvia	Phanerozoic	Genera	242	253	4	249	4
Bivalvia	Palaeozoic	Genera	198	253	13	240	9
						Total characters lost	13
Protobranchia	Phanerozoic	Specimens	40	253	126	127	126
Protobranchia	Phanerozoic	Genera	30	253	129	124	3
Protobranchia	Palaeozoic	Genera	29	253	132	121	3
						Total characters lost	126
Heteroconchia	Phanerozoic	Specimens	161	253	89	164	89
Heteroconchia	Phanerozoic	Genera	102	253	91	162	2
Heteroconchia	Palaeozoic	Genera	71	253	101	152	10
						Total characters lost	101
Pteriomorphia	Phanerozoic	Specimens	145	253	46	207	46
Pteriomorphia	Phanerozoic	Genera	107	253	47	206	1
Pteriomorphia	Palaeozoic	Genera	95	253	54	199	7
						Total characters lost	54

Table.3.0 Genera and characters retained in each data treatment (Mixed Range stratigraphy). Temporal range: the specific interval of time analysed. Condense level: the taxonomic level of taxa. Specimen count: the number of taxa after condensing. Character count: the total number of characters. Invariable pruned: the number of invariable characters removed. Clade specific remain: those characters that remain which show variability in characters states for at least one taxon. Net change: the number of characters lost with each change in the data. Number of characters lost: total number of characters lost.

Clade	Temporal Range	Condense Level	Specimen Count	Total Character Count	Characters Pruned	Remaining Characters	Net loss
Bivalvia	Phanerozoic	Specimens	350	253	-	-	-
Bivalvia	Phanerozoic	Genera	242	253	4	249	4
Bivalvia	Palaeozoic	Genera	185	253	20	233	16
						Total characters lost	20
Protobranchia	Phanerozoic	Specimens	40	253	126	127	127
Protobranchia	Phanerozoic	Genera	30	253	128	125	2
Protobranchia	Palaeozoic	Genera	28	253	139	114	10
						Total characters lost	139
Heteroconchia	Phanerozoic	Specimens	161	253	89	164	89
Heteroconchia	Phanerozoic	Genera	102	253	90	163	1
Heteroconchia	Palaeozoic	Genera	67	253	102	151	12
						Total characters lost	102
Pteriormorphia	Phanerozoic	Specimens	145	253	46	207	46
Pteriormorphia	Phanerozoic	Genera	107	253	46	207	0
Pteriormorphia	Palaeozoic	Genera	87	253	61	192	15
						Total characters lost	61

Table.3.1 Genera and characters retained in each data treatment (Treatise stratigraphy). Temporal range: the specific interval of time analysed. Condense level: the taxonomic level of taxa. Specimen count: the number of taxa after condensing. Character count: the total number of characters. Invariable pruned: the number of invariable characters removed. Clade specific remain: those characters that remain which show variability in characters states for at least one taxon. Net change: the number of characters lost with each change in the data. Number of characters lost: total number of characters lost.

diversity of bivalves in the Palaeozoic. The data was also split into the three individual subclades to create a total of four treatments of the data.

The first treatment includes all bivalve genera across the Palaeozoic including the three Cambrian bivalves (*Fordilla*, *Pojetaia*, *Tuarangia*) present in the original analysis (Carter *et al.* 2006). The next three treatments split genera into the three major subclades: the Protobranchia, Heteroconchia, and Pteriomorphia. There is the question of whether splitting the data by taxonomic grouping to conduct disparity analysis is a circular argument as some of the characters that split up the three clades are then used to look at disparity. I believe using the data herein for disparity analysis is not circular as the three main clades also support from molecular characters (Kocot *et al.* 2011). The groupings themselves are used as a means of exploring how individual groups of bivalves contribute to the Palaeozoic bivalve disparity curve. Each subclade was assigned a colour for easy identification in the results: Bivalvia (black), Protobranchia (red), Heteroconchia (blue), and Pteriomorphia (green). Subsampling the data in this way inevitably resulted in some morphological characters becoming invariable and redundant for a given subset of genera. The breakdown of genera and characters that comprised the different partitions are found in Table 3.0 and 3.1. Despite the redundancy of characters, distance calculations were possible between all genera in all treatments of the data.

### **3.6 Stratigraphic Resolution**

After taxa were amalgamated to the rank of genus, stratigraphic ranges were assigned to reflect that ranking. I used the *International Stratigraphic Chart 2009* (Ogg *et al.* 2008; Gradstein *et al.* 2004) offering a maximum resolution of 100 stages. First and last occurrences were determined at the maximum possible precision for each genus and in general, coding the maximum duration for a given genus. Taxa were assumed to persist uninterrupted between these dates. The stratigraphic data defining geological ranges of genera were assembled from several sources. Generic ranges were compiled from the Paleobiology Database. Several downloads were made, but the final version of dates came from the database on the 13<sup>th</sup> of March, 2012 (Alroy *et al.* 2001; Kiessling *et al.* 2012).

Further data was compiled from the database of Krug *et al.* (2007), character 253 of the Carter *et al.* morphology matrix (2000; 2006), and individual literature sources (Johnston & Goodbody 1989; Cope 1996; Liljedahl 1990; Schneider 1999; Perry & Larsen 2004; Sartori & Harper 2009). Diversity curves (traditional counts and shareholder quorum subsampling) were generated directly from the websites online facilities on the 10<sup>th</sup> of May, 2013. Further data on paleolatitude, stratigraphic formation, collection number, life habit and locomotion were downloaded in addition to the generic stratigraphic ranges for bivalves from the Paleobiology Database on the 28<sup>th</sup> of May, 2013 (Alroy *et al.* 2001; Kiessling *et al.* 2013).

A separate list of generic ranges was compiled from an *in prep* revision of the Treatise of Invertebrate Palaeontology, possible as a result of the collaboration between Martin Hughes (University of Bath), Dr Matthew Wills (University of Bath), and Dr Joseph Carter (University of North Carolina). The final iteration of Treatise dates used in all relevant analyses was compiled on the 27<sup>th</sup> June, 2013.

Stratigraphic ranges for each genus were formatted as binary matrices of absences and presences. First and last dates derived from the Treatise were converted to stages based on the *International Stratigraphic Chart 2009*. All results are reported at the stratigraphic resolution of stage, although longer time intervals bins (epoch and period) were explored to assess their effect on the results. In practice lower stratigraphic resolutions had no effect upon the conclusions reported throughout. By amalgamating all sources of stratigraphic data two stratigraphic treatments were possible. The 'Mixed Range' treatment resulted from amalgamating ranges from the Paleobiology Database, the database of Krug *et al.* (2007) and individual literature sources (Carter *et al.* 2006; Sartori & Harper 2009; Perry & Larsen 2004; Schneider 1999; Cope 1996; Johnston & Goodbody 1989). The treatment was an updated version of the Paleobiology Database data, revising outdated ranges and filling in gaps (e.g. *Eodon*).

The 'Treatise' treatment was composed of the stratigraphic ranges mined from the *in prep* edition of Treatise of Invertebrate Palaeontology. The dates making up the Treatise treatment have been compiled whenever possible in isolation of the Paleobiology Database (personal communication Dr Joseph Carter) and can therefore be considered a separate source for stratigraphic data.

### **3.7 Disparity Specific Methods**

#### **3.8 Use of Discrete Rather than Continuous Characters**

Discrete characters were chosen over continuous characters for several reasons. Bivalves share a common and deceptively restrictive body structure i.e. the soft body (mantle) contained within the two valves. Despite this, they have evolved considerable morphological variation in discrete hard and soft characters, including but not limited to: the absence/presence/shape of auricles (hinge "wings"), absence/presence of surface structures such as crenulations or spines, shell symmetry/asymmetry, muscle scar shape, ligament/hinge/tooth shape, shell aperture size, composition of the shell, size/structure of the siphon, size of foot, presence/absence of brood poach, presence/absence of photoreceptors on the mantle edge. These characters have been used to recognised over 9000 extant species (Huber 2010), with thousands more cited in the literature e.g. 26583 unique species instances are catalogued in the Paleobiology Database (as of the 04/10/2012). From this rich body of literature and diversity, there is a wealth of information available that can be used to create discrete character datasets that can build up empirical morphospaces of bivalves.

Although landmark, outline, or continuous measurement data are alternatives to discrete characters. Traditionally these kinds of data have been used to infer the subtle changes in shape between small numbers of taxa. For example the study of Schneider *et al.* (2010) collected 1246 outlines across three genera, while the study of Grey *et al.* (2010) collected 1855 measurements for a single genus. As the number of genera in this section number over 200, collecting the same numbers of measurements for each would have been prohibitively time consuming and would require access to primary specimens which was not possible. The degree of morphological convergence between taxa is also a problem when using continuous characters.

This could be equally true for discrete characters. With the construction of a phylogenetic tree homoplastic characters could be pruned. The tree built upon the morphological characters of the Carter *et al.* dataset (2000; 2006) was not available at the time of analysis and so could not be used in this capacity. A direction of future work could be to take the phylogenetic tree of bivalves built using the Carter matrix and constructing a disparity dataset built upon an additional set of measurement or outline/landmark data. The disparity within and between clades along the tree could then be assessed as in other disparity studies (Slater *et al.* 2010; Burbrink & Pyron 2010; Dornburg *et al.* 2011; Smith *et al.* 2011).

The data herein are not entirely devoid of continuous data. Some discrete character states such as the size of the auricle includes continuous variation split into a discrete characters, while states of the teeth such as taxodonty or heterodonty (discontinuous in modern taxa) follow more continuous integration in the fossil record. This does not affect the disparity analyses herein as all variables in the morphology have been treated as discontinuous.

Collection and analysis of landmarks, outlines and measurement data could be the grounds for future projects, including the quantification of allometric disparity trajectories of bivalve clades. This would allow us to explore the role that developmental processes have on the evolution in of body form in bivalves. One notes that in other studies where disparity has been estimated using both continuous and discrete characters for the same sets of taxa, the results can be in marked agreement (Foth *et al.* 2012).

There is also a more practical reason for choosing discrete characters over continuous character data. The script developed in Chapter 1 provided a semi-automated system that uses *R* (2012) to construct discrete character morphospaces and measures of disparity through time (Hughes *et al.* 2013). This system is immediately applicable to the matrix of Carter *et al.* (2000; 2006).

	11	25	51	55	56	61	67	107	119	144	145	146	147	148	149	150	151	204	212	213	214
A	1	10	4	5	10	10	10	10	10	10	-	9	9	9	8	9	9	10	2	10	4
B	2	11	5	6	11	11	11	-	11	11	-	10	10	10	9	10	10	-	3	11	5
C	3	12	6	7	-	-	-	-	12	12	-	11	-	11	10	11	11	-	-	12	6
D	4	13	-	8	-	-	-	-	13	13	-	12	11	12	11	12	12	-	4	13	7
E	5	14	7	-	-	-	-	-	14	-	-	13	12	13	12	13	-	-	5	14	-
F	6	15	8	9	-	-	-	-	-	14	-	14	-	-	-	-	-	-	6	15	-
G	7	16	9	10	-	-	-	-	-	15	-	15	13	-	-	-	-	-	7	-	-
H	8	17	10	11	-	-	-	-	-	16	-	-	-	-	-	-	-	-	8	-	-
I	9	-	11	12	-	-	-	-	-	17	-	16	-	-	-	-	-	-	-	-	8
J	10	18	12	13	-	-	-	-	-	18	-	-	-	-	-	-	-	-	9	-	9
K	11	19	13	14	-	-	-	-	-	19	-	-	-	-	-	-	-	-	10	16	10
L	12	-	-	-	-	-	-	-	-	20	8	-	-	-	-	-	-	-	-	17	11
M	-	-	-	15	-	-	-	-	-	21	9	-	-	-	-	-	-	-	-	18	-
N	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	19	12
O	-	-	14	16	-	-	-	-	-	22	-	-	-	-	-	-	-	-	-	20	-
P	-	20	15	17	-	-	-	-	-	-	10	-	-	-	-	-	-	-	-	21	13
Q	-	-	16	18	-	-	-	-	-	23	11	-	-	-	-	-	-	-	-	-	14
R	-	-	17	19	-	-	-	-	-	-	12	-	-	-	-	-	-	-	-	22	15
S	-	-	18	20	-	-	-	-	-	-	13	-	-	-	-	-	-	-	-	23	16
T	-	-	-	-	-	-	-	-	-	-	14	-	-	-	-	-	-	-	-	24	-

Table.3.2 Conversion table showing values given for those morphological characters (columns) of the original matrix of Carter *et al.* that use letter codes for their character states (rows). Dashes are inapplicable states for those characters.

### **3.9 Character Coding**

The original matrix used letter codes for some specific states whenever numbered states were insufficient. These were re-coded at numbered states (Table 3.2). This process was required by the software for distance calculation. No character ordering was used. This treated all characters states with values greater than one in the same manner i.e. a single unit distance from 0. As the specimens were condensed to generic level to combat uneven sampling, character states were merged in the following way. (1) Where there was no character state conflict resulting from the condensing, no action was taken. (2) If the character states conflicted between specimens, I assigned the state most frequently observed for that character for those specimens. Polymorphic character states constitute <2% of the data. As a conservative approach, they were treated as NA prior to condensing of specimens. Missing and/or gap state data were more prevalent (10% of the data). These states were also treated as NA prior to condensing. Even at this concentration, it was possible to calculate pairwise distances between all genera; even when considering the three subclades individually.

### **3.10 Intertaxon Distance Matrices and Ordination**

The above procedures resulted in a single morphological character matrix and one stratigraphic presence/absence file for each treatment. Several treatments were produced. The first treatment included all Palaeozoic bivalve genera that make up the original morphology matrix. From this matrix, the three subclades were identified and their morphological and stratigraphic data saved as separate files: Protobranchia only, Heteroconchia only and Pteriomorphia only. All subsequent analyses were conducted using *R* scripts originally published in Hughes *et al.* (2013). I first generated a generalised Euclidean Distance (GED) matrix (Wills 1998b) between all pairs of genera. For each pair of genera, the difference between each comparable character state was squared, these differences summed, divided by the total number of comparable states, then square rooted (Wills 1998b). Distance matrices were ordinated using principal coordinates analysis (PCO) (Wills *et al.* 1994). Missing character states result in negative Eigen values. To reduce these, I implemented Cailliez's correction (Cailliez 1983).

Some published studies chose to calculate disparity on a subsample PCO axes, selecting those axes that capture 90% of the variance in the data (Foote 1999; Wills *et al.* 2012). When quantifying disparity (herein the sums of variances), all PCO axes were retained as the majority (85%) of PCO axes were required to explain at least 90% of the variance. The standard procedure of using a subset of PCO axes has involved uncorrected PCO, i.e., PCO without Cailliez's correction (Foote 1999). The recent studies of Wills *et al.* (2012) instead selected a proportion of Cailliez corrected axes. The effect that Cailliez's correction has on the choice of PCO axes was explored. Using the Bivalvia

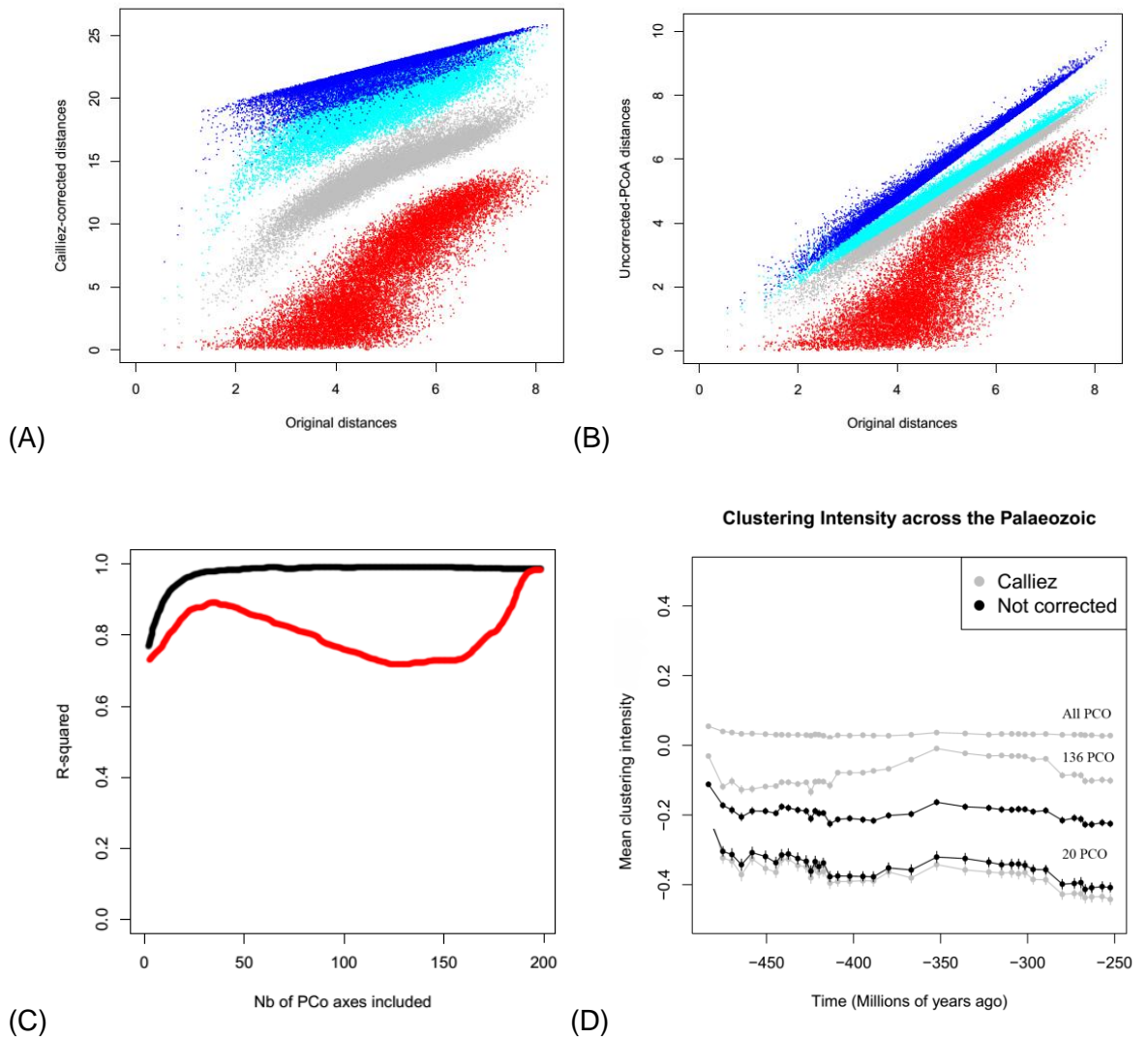


Figure.3.3: PCO distances vs. original distances (character dissimilarities) of bivalve dataset (A) Cailliez corrected PCO distances; (B) uncorrected PCO. Red, grey, light blue and dark blue correspond to 2, 40, 125 and 180 axes included respectively. (C) R squared distance between PCO distances and original distances vs. the number of PCO axes included. Black line is uncorrected; red is Cailliez corrected. (D) Clustering intensity through time for bivalves as a whole using critical values of PCO axes: 20, 136 and all axes. Black are uncorrected PCO values and grey Cailliez corrected. Error bars represent standard error values calculated from 20 replicates.

dataset, Cailliez's correct distances (PCO) and non corrected distances were compared to original distances (character dissimilarities) (Fig.3.3).

The effect of Cailliez's correction on the distances was evident in Fig.3.3A. The different colours represent different numbers of PCO axes selected: 2, 40, 125 and 180 axes for red, grey, light blue and dark blue respectively. The curvature of the different distributions of points in Fig.3.3A shows that although the ranking of distances among taxa remains the same, the relationship between original and corrected distances is not linear. As the number of axes increase, the curved left tail becomes less densely populated. When all axes are used the relationship is a straight line. The uncorrected



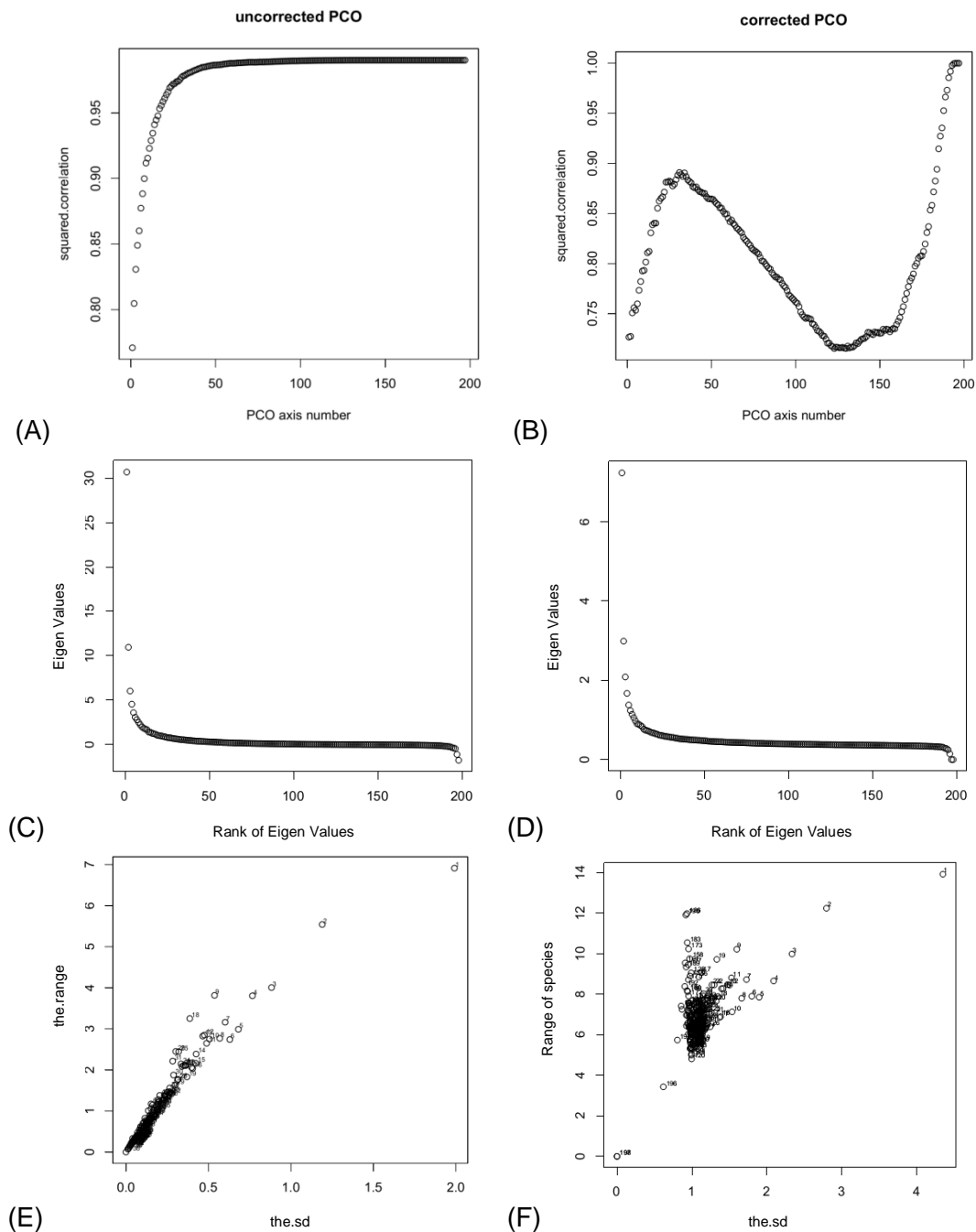


Figure.3.4: Choosing number of PCO axes for clustering analysis of Palaeozoic bivalves. (A) Squared correlation between morphological distance matrix (dissimilarity) and sequential addition of principal coordinate axis uncorrected for negative Eigen values. (B) Same using Cailliez corrected PCO. (C) Variation explained by each PCO axis (uncorrected). (D) Variation explained by each PCO axis (corrected). (E) Standard deviation vs. range of species scores on uncorrected PCO. (F) Standard deviation vs. range of species scores on corrected PCO. Each point represents an individual PCO axis.

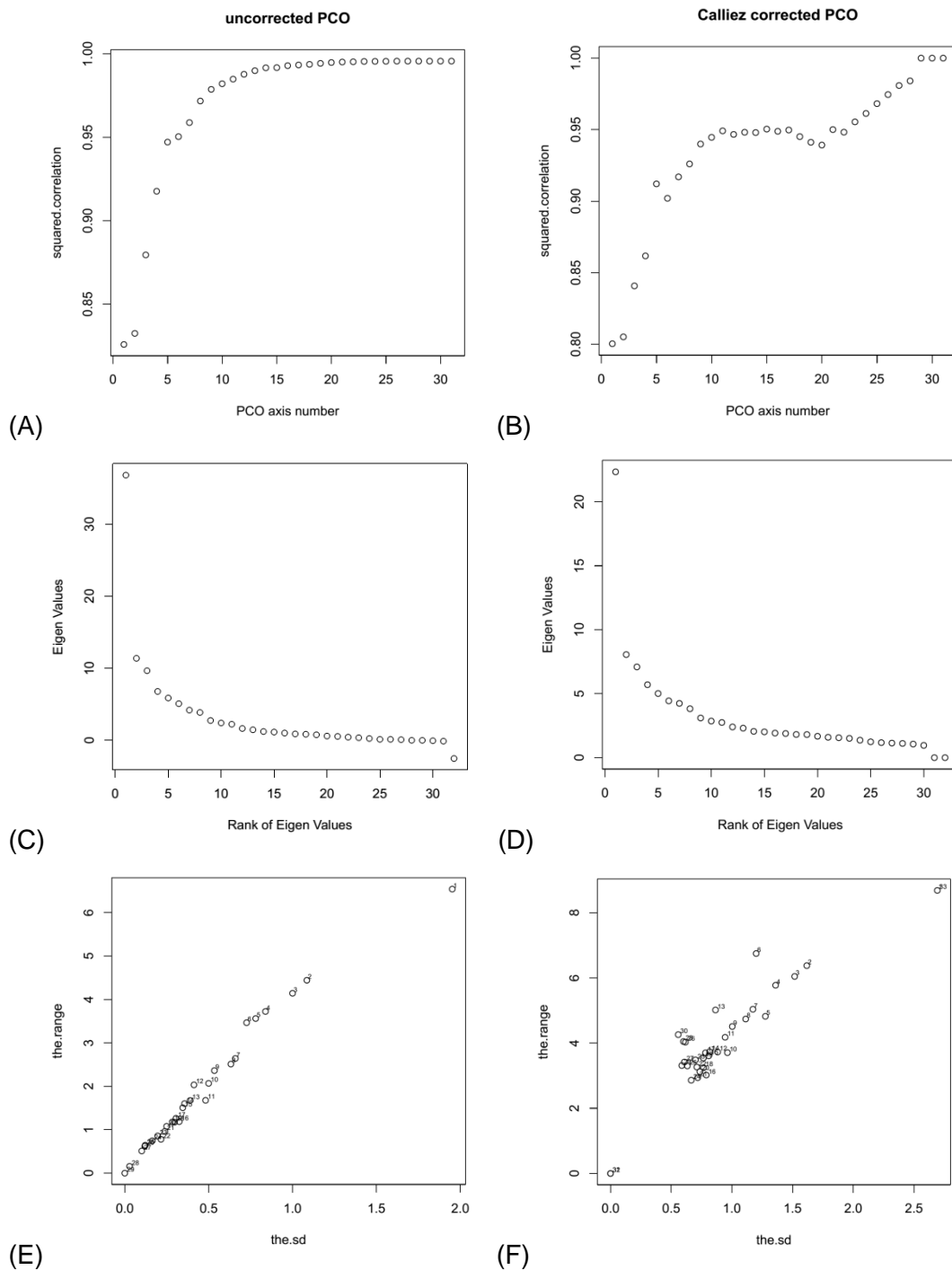


Figure.3.5: Choosing number of PCO axes for clustering analysis for Protobranchia. (A) Squared correlation between morphological distance matrix (dissimilarity) and sequential addition of principal coordinate axis uncorrected for negative Eigen values. (B) Cailliez corrected PCO. (C) Variation explained by each PCO axis (uncorrected). (D) Variation explained by each PCO axis (corrected). (E) Standard deviation vs. range of species scores on uncorrected PCO. (F) Standard deviation vs. range of species scores on corrected PCO. Each point represents an individual PCO axis.

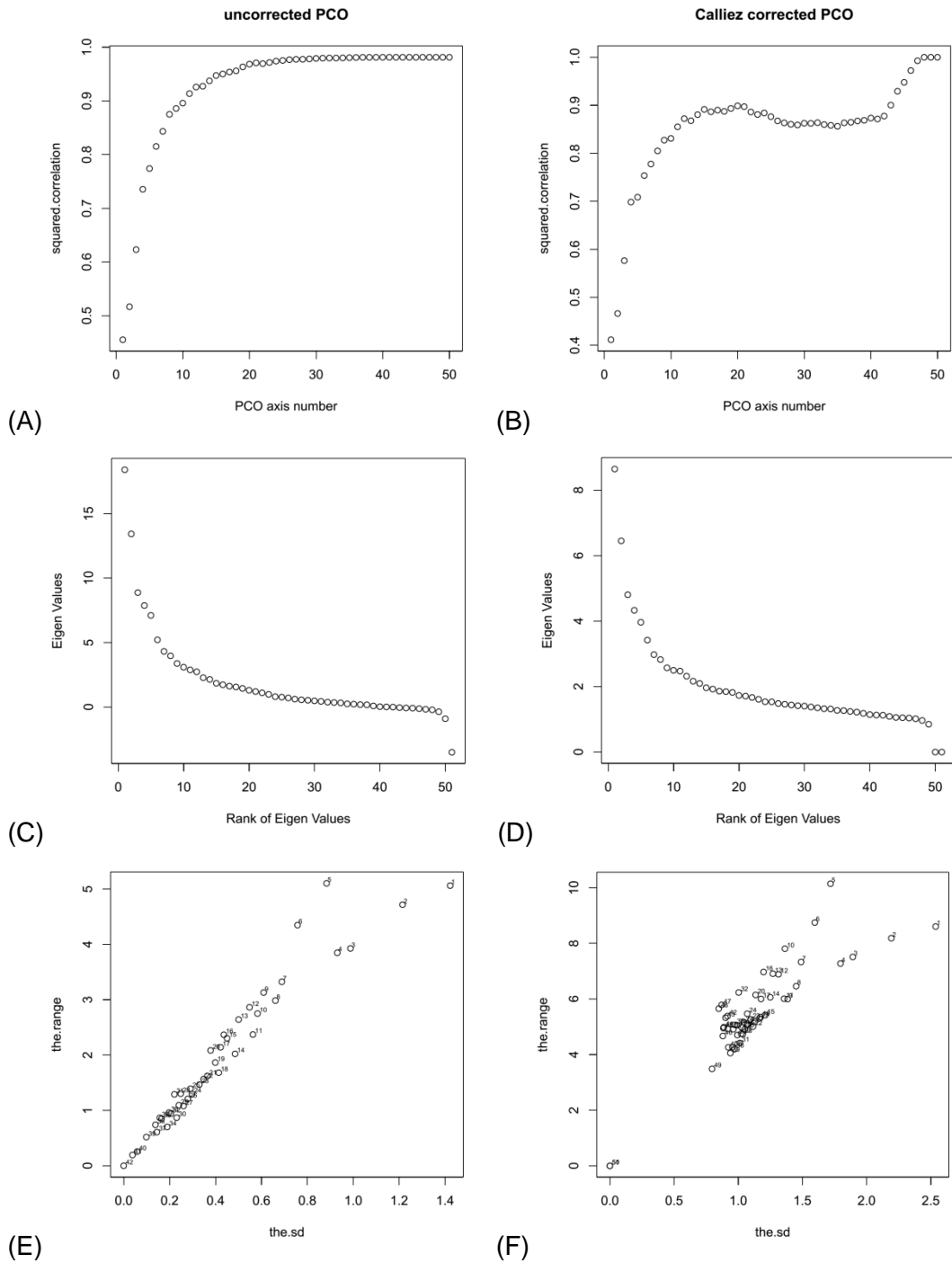


Figure.3.6: Choosing number of PCO axes for clustering analysis for Heteroconchia. (A) Squared correlation between morphological distance matrix (dissimilarity) and sequential addition of principal coordinate axis uncorrected for negative Eigen values. (B) Cailliez corrected PCO. (C) Variation explained by each PCO axis (uncorrected). (D) Variation explained by each PCO axis (corrected). (E) Standard deviation vs. range of species scores on uncorrected PCO. (F) Standard deviation vs. range of species scores on corrected PCO. Each point represents an individual PCO axis.

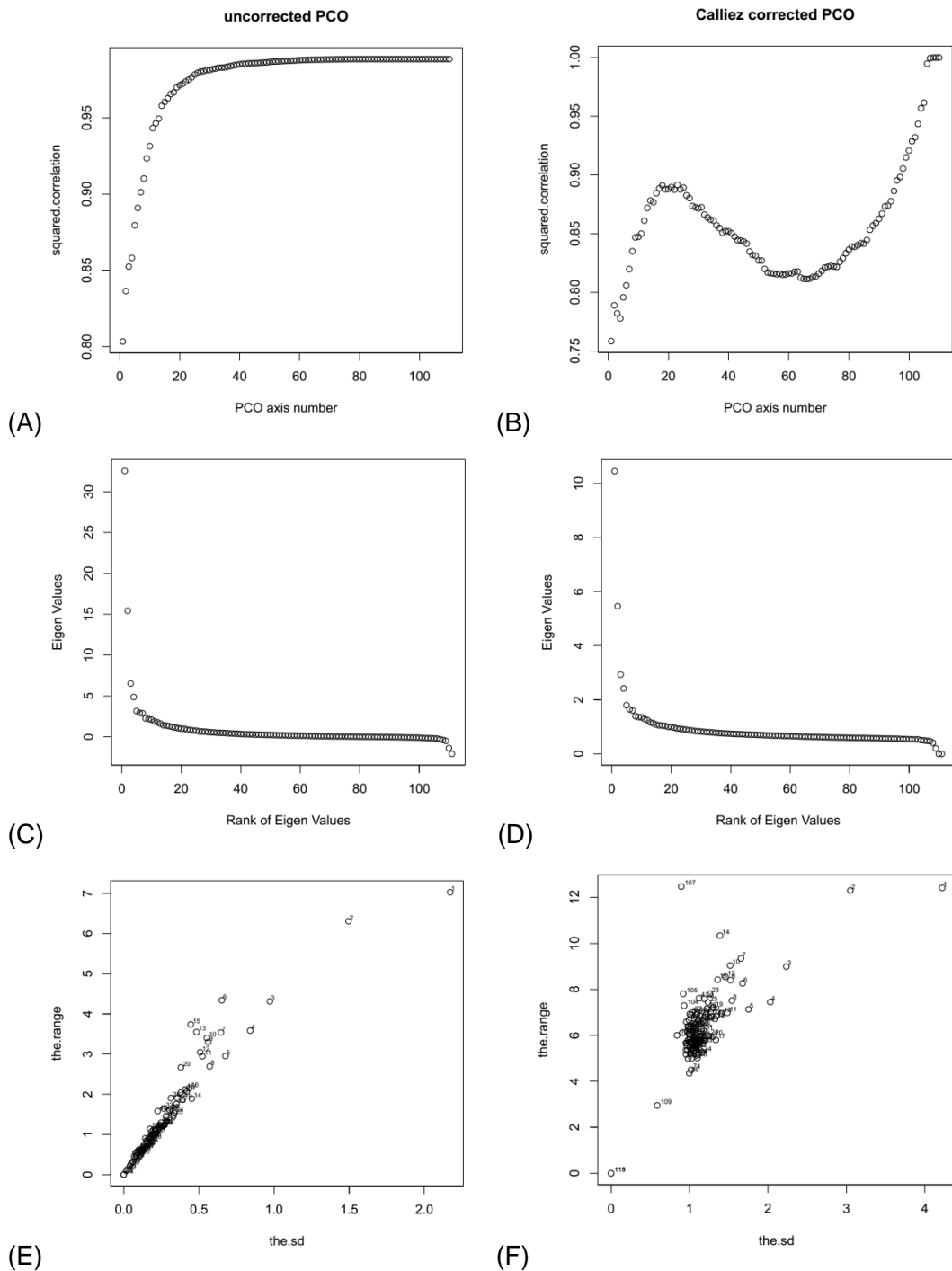


Figure.3.7: Choosing number of PCO axes for clustering analysis for Pteriomorphia. (A) Squared correlation between morphological distance matrix (dissimilarity) and sequential addition of principal coordinate axis uncorrected for negative Eigen values. (B) Cailliez corrected PCO. (C) Variation explained by each PCO axis (uncorrected). (D) Variation explained by each PCO axis (corrected). (E) Standard deviation vs. range of species scores on uncorrected PCO. (F) Standard deviation vs. range of species scores on corrected PCO. Each point represents an individual PCO axis.

PCO in Figure 3.3B does not show this non linearity. The  $r^2$  squared distances between the corrected PCO axes and uncorrected PCO axes are plotted in Figure 3.3C using the method of Foote (1999). It shows the percentage of the data explained as axes are added. The black line shows the uncorrected PCO while the red is the Cailliez's corrected PCO. The result of Figure 3.3A-C shows that taking all axes into account for disparity values provides the same results whether Cailliez or uncorrected PCO is used. The effect on clustering (Fig. 3.3D) increased as the number of axes increased and whether Cailliez correction was used. Therefore the number of PCO axes were carefully chosen for the cluster analysis of Bivalvia and subclades. Future work should explore the effect of Cailliez correction on disparity analyses to test whether the non linear properties found in this brief exploration are unique to the bivalve dataset or of wider importance. Cailliez correction on disparity analyses to test whether the non linear properties found in this brief exploration are unique to the bivalve dataset or of wider importance.

The results of Figure 3.3A-D show that the number of axes retained can be important for the interpretation of the degree of clustering to avoid projection effects that mimic densely populated regions of morphospace. As the number of axes is increased the pattern of clustering intensity through time (Fig. 3.3D) became less negative and less variable through time. The use of corrected versus uncorrected PCO does not seem to have a great impact when the number of axes were low. Therefore the number of axes for cluster analysis of the four clades: Bivalvia, Protobranchia, Heteroconchia and Pteriomorphia were determined using three methods. First method was to find the squared correlation between the raw squared distances among the bivalve genera and squared distances based on the PCO (uncorrected and corrected) (Foote 1999). The results can be viewed in Figures 3.4-7AB.

The plots show the  $r^2$  value between the raw distances between taxa (character dissimilarities) and the distances based on increasing numbers of principal coordinate axes. As principal coordinate axes increased, the  $r^2$  value, i.e. the amount of variation in the distance matrix explained by the number of axes, should increase. This is the expectation as seen using the uncorrected PCO axes (Fig. 3.4-7A), but as shown in Figure 3.3C, not in the Cailliez corrected. The Cailliez corrected  $r^2$  result suggests a non linear relationship between the raw distances and the Cailliez corrected PCO for values between 40-190 axes. In Figures 3.4-7AB, the plots show the distribution of raw distances vs. the corrected/uncorrected PCO distances for critical numbers of PCO axes. The left hand curved tail seen in the corrected distances (Fig. 3.4-7B) again suggests this non linear relationship. As corrected and uncorrected distances up to 20 axes provide similar results, 20 axes were used for calculating clustering in the Palaeozoic Bivalvia treatment.

The squared correlation is the same as looking at the amount of variation explained by each PCO axes (Fig. 3.4-7CD). These results show no differences between the corrected and uncorrected treatments. The final method looked at the variance

structure of each PCO axes by plotting the standard deviation and range of PCO coordinates (Fig. 3.4-7EF) (Foote 1999). If the variance structure was the same along each PCO axis, differing only in scale then Figures 3.4-7EF should show a linear correlation. It can be seen that for the corrected PCO data (Fig. 3.4-7BF) between PCO axes 20-190 that the variance structure is different for a number of PCO axes that share similar standard deviations but a wider range. I decided that a subset of axes would be used for calculating clustering as it can be sensitive to the number of axes used (Fig.3.3D). It is shown that for the first 20 axes corrected and uncorrected patterns converge, giving similar results (Fig.3.3D). Therefore for calculating levels of clustering in the Bivalvia, 20 axes selected. For the subclades (Fig.3.5-7), the plots point to using 10, 15 and 20 axes for Protobranchia, Heteroconchia and Pteriomorphia respectively.

The resultant ordination can be visualised in two or three dimensions as a morphological space (morphospace). Each point represents a genus, and each axis describes changes in shape. To visualise the exploration of genera through the morphospace, the first three principal coordinates were selected. Since scores on all coordinates were used to calculate disparity (as the sum of the univariate variances) and clustering (squared mean sample deviation and mean proportion difference of distances), the ordination step was primarily for visualisation purposes. Non-metric multidimensional scaling (NMDS) (Anderson *et al.* 2011; Shen *et al.* 2008) was not chosen because I wished to retain the metric properties of the distance matrices (Huntley *et al.* 2006).

### **3.11 Quantifying Clustering**

To measure changing levels of clustering in morphospace through time, two methods were used. One of the two methods was the principal points method of Flury (1990; 1993) and Klingenberg & Froese (1991). The aim of the method was to assess the degree of taxon clustering in morphospace compared to that of a multivariate uniform distribution constrained to have the same spatial bounds as the empirical distribution. The observed distribution was then plotted alongside the null expectation that the dispersal of points in the empirical distribution was random and uniform. To measure clustering of the empirical distribution the method uses a sequentially increasing series of  $k$  points sampled from the empirical distribution. These principal points minimise the Euclidean distance between any one point and its nearest neighbour. As more points are added to the morphospace, the ability of  $k$  points to accurately approximate the empirical distribution increases.

To assess how well the principal points approximate the empirical distribution,  $k$  points (from 1 to  $N$ , the number of observations) were assigned to the space (200 iterations); the  $k$  mean was taken and the sample mean squared deviation (SMSD) was calculated (Wills *et al.* 2012). The same procedure was done for  $k$  points taken from a multivariate uniform distribution. The SMSD can range from the variance of the sample

(multiplied by  $(N-1)/N$ ) down to zero. As  $k$  ranged from 1- $N$ , SMSD decayed. When compared with results derived from the uniform distribution, any clustered SMSD values that are outside the range of the 95% confidence limits of the uniform distribution are said to be evidence of significant clustering (Wills *et al.* 2012). The method was used to assess clustering of the space as a whole (i.e. the Palaeozoic morphospace), and for each period of the Palaeozoic.

The other method utilised the nearest neighbour analysis of Foote (1990). This also measured clustering of the space but in a different way to principal points analysis. The method calculated the smallest distance ( $d$ ) between each taxon ( $i$ ) with all other taxa. It then compared these observed distances ( $d_i$ ) with those expected ( $r_i$ ) from a null model of a random uniform distribution that has the same constraints (same number of neighbours and spatial bounds) as the empirical one. For each  $d_i$  and  $r_i$ , the proportional difference was obtained ( $p_i = (d_i - r_i) / r_i$ ) and the mean  $p_i$  calculated (Foote 1990; Wills *et al.* 2012). The mean value was used as the intensity of clustering. Values of  $p_i < 0$  indicate clustering while  $p_i > 0$  suggests points were further away than expected. Values of  $p_i \approx 0$  suggest no clustering and instead suggest taxa were randomly distributed. Again clustering was calculated for the datasets: Bivalvia, Protobranchia, Heteroconchia and Pteriomorpha.

### **3.12 Quantifying Disparity & Partial Disparity**

Disparity was quantified using the sum of univariate variances index (Foote 1991a), which measured the spread of genera through morphospace. The disparity values were calculated using part of the semi-automated R script designed in Chapter 1. Genera in each Stage were isolated, their principal coordinate scores were identified, the covariance matrix calculated and the diagonal of the covariance's summed. Taxa were then sampled with replacement (Efron 1982) 1000 times. The producer estimated mean and standard error of disparity in each stage. The sum of variances index was favored because of its relative insensitivity to sample size differences and its complete indifference to the orientation of the coordinate axes upon which it is being calculated. The first property was useful because the number of taxa sampled can be highly variable through time, and because larger samples will tend to have a wider distribution within their character space. Uncorrected measurements of range are sensitive to these changes in sample size, to the extent that they may largely reflect proxies of diversity.

The second property was useful because although the orientation of the ordination is derived from all taxa simultaneously, taxa within a given time bin are almost invariably some subset of these. Were these subsets to have been ordinated independently, the orientation of axes would very likely have been different. Disparity was calculated for all time intervals over which each clade persisted. In order to determine the contribution of

subgroups ( $j$ ) to the total morphospace in each interval, I used the method of Foote (1993a) to calculate partial disparity (PD).

$$PD_i = \frac{\sum_j d_{ij}^2}{(N - 1)} \quad (1)$$

The morphospace coordinates of genera ( $j$ ) in each stage were isolated. The sum of squared distances of the  $j$ th point in subgroup  $i$  from the overall centroid of the total morphospace was divided by the total number of points ( $N$ ) minus one (equation 1). The final result was the contribution of each subclade to the total level of disparity in each interval.

### 3.13 Testing for Trends

Several methods were employed to quantify the magnitude and statistical significance of apparent disparity trends across the Palaeozoic. Individual disparity trajectories (subclades combined and separate) used Stage, Epoch or Period interval binning for the Mixed Range and Treatise treatments. For each trajectory I calculated Spearman Rank correlation (and associated p-values) between disparity and time. The same techniques were applied to the first differences of disparity, i.e., the differences in disparity between successive time intervals. This was used as complementary technique to likelihood ratio tests (Friedman 2010) that identify jumps in disparity along the time series. To identify whether these trends represent directional, random walk, static or Ornstein–Uhlenbeck processes of morphological evolution I applied the *R* package *paleoTS* of Hunt (Hunt 2006; Hunt *et al.* 2008; Hunt 2008).

These models have traditionally been used to test whether lineages have evolved in a random or adaptive manner (Geary *et al.* 2010), but have also been used to show individual traits can each conform to specific evolutionary models (Hopkins & Lidgard 2012). A disparity trajectory (as it is calculated here) is the summative effect of variation in many characters across numerous genera interpolated along the evolutionary history of a clade. Therefore the results of these analyses correspond to the evolutionary trajectory of the clade rather than any individual character.

In this case the modeling approach (Hunt 2006) used maximum likelihood to test whether the disparity of bivalves followed trends most consistent with evolutionary models for generalised random walk (GRW), unbiased random walk (URW), stasis (STA) or Ornstein–Uhlenbeck (OU) processes. For each disparity trajectory, the method uses the disparity values through time, applying functions that find the maximum likelihood solutions to the data based on the evolutionary models. Time was modeled as discrete increments. Each increment involved an evolutionary change (i.e. disparity increased, decreased or stayed the same) drawn at random from a distribution of evolutionary



transitions (steps) (Hunt 2008; Hunt 2006). The long term changes are dependent on the mean ( $\mu_s$ ) and variance of the step distribution ( $\sigma_s^2$ ) (Hunt 2008; Hunt 2006).

The major difference between the GRW and URW models is the direction of the random walk, which is effected by the  $\mu_s$ . When this value is zero, there is no directionality and it becomes similar to the URW model. When the mean increases or decreases the result is a GRW in the positive or negative direction. For stasis, the mean is fixed (with a level of variance  $\omega$  associated with that value) around some phenotypic optimum ( $\theta$ ). It is different from the URW as there is no net accumulation of change with time. The Ornstein–Uhlenbeck process is effectively a mixture of the GRW and stasis models. The model begins as highly directionalised, followed by a slowing down and tapering off towards some phenotypic optimum (Hunt 2008).

Two different parameterisations are possible for calculating the means and step variance values: ancestor-descendent (AD), and joint. They are needed to adequately fit the models to the data as the models themselves make predictions as to how the data is distributed. Ancestor-descendent parameterisation considered the morphological differences in each ancestor-descendent pair as separate populations (Hunt 2006). Joint parameterisation weighs the joint distribution of trait values across all sample populations simultaneously (Hunt 2008). Essentially the former parameterisation (AD) the models use the normal distribution while the latter (joint) makes the models use a multivariate normal distribution. The current *PaleoTS* script can only use AD and joint parameterisations on the first three models: GRW, URW and STA. The OU model only implemented joint parameterisation. When testing the three models GRW, URW and STA on their own both AD and joint parameterisations were used (but gave the same answers).

The equations for calculating expected values, variances and covariance's for each of the evolutionary models are set out below:

#### General random walk (GRW)

$$\text{Expected: } E[X_i] = X_0 + \mu_s t_i \quad (2)$$

$$\text{Variance: } Var[X_i] = \sigma_s^2 t_i + \varepsilon_i \quad (3)$$

$$\text{Covariance: } Cov[X_i, X_j] = \sigma_s^2 \cdot \min(t_i, t_j) \quad (4)$$

#### Unbiased random walk (URW)

$$\text{Expected: } E[X_i] = X_0 \quad (5)$$

$$\text{Variance: } Var[X_i] = \sigma_s^2 t_i + \varepsilon_i \quad (6)$$

$$\text{Covariance: } Cov[X_i, X_j] = \sigma_s^2 \cdot \min(t_i, t_j) \quad (7)$$

#### Stasis (STA)

$$\text{Expected: } E[X_i] = \theta \quad (8)$$

$$\text{Variance: } Var[X_i] = \omega + \varepsilon_i \quad (9)$$

$$\text{Covariance: } Cov[X_i, X_j] = 0 \quad (10)$$

Ornstein–Uhlenbeck (OU)

$$\text{Expected: } E[X_i] = [1 - \exp(-\alpha t_i)]\theta + \exp(-\alpha t_i)z_0 \quad (11)$$

$$\text{Variance: } \text{Var}[X_i] = (\sigma_s^2/2 \propto)[1 - \exp(-2 \propto t_i)] + \varepsilon_i \quad (12)$$

$$\text{Covariance: } \text{Cov}[X_i, X_j] = \left(\frac{\sigma_s^2}{2} \propto\right) \exp(-\propto t_{ij}) [1 - \exp(-2 \propto t_{min})] \quad (13)$$

Where  $\mu_s$  is the mean and  $\sigma_s^2$  is the step variance,  $X$  is the trait (disparity in this case),  $X_0$  is the trait at the beginning of the sequence,  $\theta$  is the phenotypic optimum (defined as the last trait value in the sequence),  $\omega$  the variance around the optimum,  $\propto$  is the restraining force around the optimum,  $t_i$  is the time from the beginning of the time sequence to the sample  $i$ ,  $t_{ij}$  is the time separating samples  $i$  and  $j$ , and  $\varepsilon_i$  is the magnitude of the sampling variance. Min is the minimum (Hunt 2006; Hunt 2008).

The four models differ in complexity, and as more complex models have the property of having a higher likelihood score due to this complexity, the AICc and Akaike weights (produced as part of the output from *paleoTS*) were used as a means of distinguishing which model most appropriately fit the data (Anderson *et al.* 2000). The AICc score is a corrected version of the AIC score which is a metric balancing the goodness-of-fit of the model vs. model complexity. The corrected version was deemed more appropriate as it is better able to handle data of modest sample sizes (Hunt 2006; Geary *et al.* 2010). The lower the AICc score, the better the model fits to the data. There is still some subjectivity in choosing a model as the AICc scores of different models can be numerically close. Akaike weights convert these AICs scores into values that sum up to 1 across the different models under scrutiny, providing an easy means to judge which model best explains the data. (Anderson *et al.* 2000).

The GRW, URW, STA and OU models have been used for invoking explanations of directional selection, neutral genetic drift and stasis and adaptive radiations (Butler & King 2004; Hunt 2008; Geary *et al.* 2010). The uses of these models have never been explored in disparity analyses. In the study of Hunt (Hunt 2008), the author commented that for the STA and GRW/URW models, the trajectory of the clade through its morphospace would be different i.e. clades following the STA model would remain stationary, while URW and GRW clades would drift in the space. By extension of his logic, OU models would be expected to show a clades drifting towards some point in space remain stationary thereafter. For disparity analyses these models can act as another descriptor of clade trajectories through the space.

### 3.14 Quantifying Significant Jumps in Disparity Through Time

As well as finding overriding trends in disparity through time, testing for shifts in disparity across important intervals such as mass extinction events can provide evidence for the sensitivity of clades to those events (Friedman 2010). I used the marginal likelihood ratio method described in Friedman (2010) hitherto used to calculate the jumps

in disparity between successive time intervals in the late Cretaceous / Early Palaeogene. I have used the same procedure as Finarelli and Flynn (2007) for the entire Palaeozoic using different time bin resolutions: Stage, Epoch and Period. I also chose to test for significant jumps in disparity of the Ordovician (Hirnantian – Rhuddanian) and Devonian (Frasnian - Famennian) mass extinctions. I did this by isolating those genera that became extinct at the boundary and those that survived the event. I calculated their disparity as per the method previously described and then the calculated the likelihood ratio across the boundary based on the disparity values.

The likelihood ratio method is applicable because I am using variance as my measure of disparity (Friedman 2010). It operates by isolating the disparity values in paired successive time intervals and calculating the likelihood ratios between null and alternative hypotheses. The null hypothesis ( $H_0$ ) is that the variance found in that interval is equal to the sample variance. The alternative hypothesis ( $H_1$ ) is that the variance in the interval is equal to the sample variance of the last interval. The equation is given below:

$$L \propto \sigma^2 \left( \frac{-(n-1)}{2} \right) \times \exp \left( \frac{-(n-1)s^2}{2\sigma^2} \right) \quad (14)$$

The marginal likelihood is  $L$ , the estimated disparity in a bin is  $\sigma^2$ , the actual variance in the bin is  $s^2$  and  $n$  is the sample size (unique number of genera) in the bin. The likelihood is calculated for each hypothesis and the ratio ( $H_0 / H_1$ ) calculated. Ratio values above eight are taken as evidence that the variance in the interval in question is significantly different from the variance of the preceding interval (Royall 1997).

### **3.15 Results**

#### **3.16 Diversity Across the Palaeozoic**

The diversity of bivalves across the Phanerozoic (and with it the Palaeozoic) was addressed in Chapter 2. The diversity results of this chapter highlight the main differences in the diversity drawn from the Mixed Range and Treatise treatments of the Carter *et al.* (2000; 2006) data. As the aim of Chapter 2 was to assess the changes in diversity at the Phanerozoic level, some of the details of the Palaeozoic specific bivalve diversity record are shown here in more detail. Specifically the results here are from the Paleobiology Database as the data can be used to show the traditional counts and SQS results for the Bivalvia and its constituent subclades, and also analyse diversity using the sampling proxy modeling. The techniques of Chapter 2 (SQS, sampling proxy modeling and range uncertainty) were also applied to the Carter *et al.* data which can be found in Appendix.2. These extra analyses were carried out to see whether the subsampled diversity would provide similar results to those using the whole bivalve fossil record. Additionally the Carter *et al.* (2000; 2006) data was used to explore the effect that filling in gaps in the

occurrence data would have on the SQS results (Appendix.2). This was done because the occurrence data of the Paleobiology Database can have gaps where fossils are missing but assumed to be present. Adding values of 1 to these gaps makes the technique treat the taxa in these intervals as present but rare.

The generic diversity of the three major bivalves subclade comprising the morphology matrix (Carter *et al.* 2000; Carter *et al.* 2006) can be found in Figures 3.8-9. It is clear from the level of diversity recorded by Figure 3.8 and 3.9 that genera contained within the morphology matrix are a small subsample of the total diversity across the Phanerozoic (Fig.2.1-3 Chapter 2). Compared to the earlier (more complete) plots, diversity in the morphology matrix is heavily biased towards the Palaeozoic, with only a few tens of genera making up the Mesozoic and Cenozoic components.

These results are caused by two factors. Firstly, the morphological data were originally collated to infer phylogeny rather than to estimate patterns of disparity. Secondly, there is a compromise between sampling all well described species within a clade (which is desirable) and the size of the clade that can be analysed (sampling all species or genera within an order is usually impracticable). The fact that the morphological data comes from a subset of the total diversity of the Palaeozoic is not problematic in itself. The selection strategy of Carter *et al.* (2000; 2006) sampled genera across as many of the Palaeozoic families as possible to infer basal bivalve relationships. Although some genera were not sampled, these came from rare families made up of poor fossil material. Therefore the sampling strategy captured to the best of its ability the majority of morphologies present in the Palaeozoic even though it did not sample all possible diversity.

The major difference between the Figures 3.8 and 3.9 is the source of the stratigraphic ranges. The data of Figure 3.8 comes from the Mixed Range treatment (essentially an updated version of the Paleobiology Database). Figure 3.9 is made up of the ranges from the Treatise treatment. Prior to the end Permian there are some differences in the distribution of taxa through the periods. The Treatise treatment had more contains more bivalve genera in the Ordovician than Silurian (Fig.3.9 black line), while in the Mixed Range treatment the opposite is true (Fig.3.8 black line). This comes

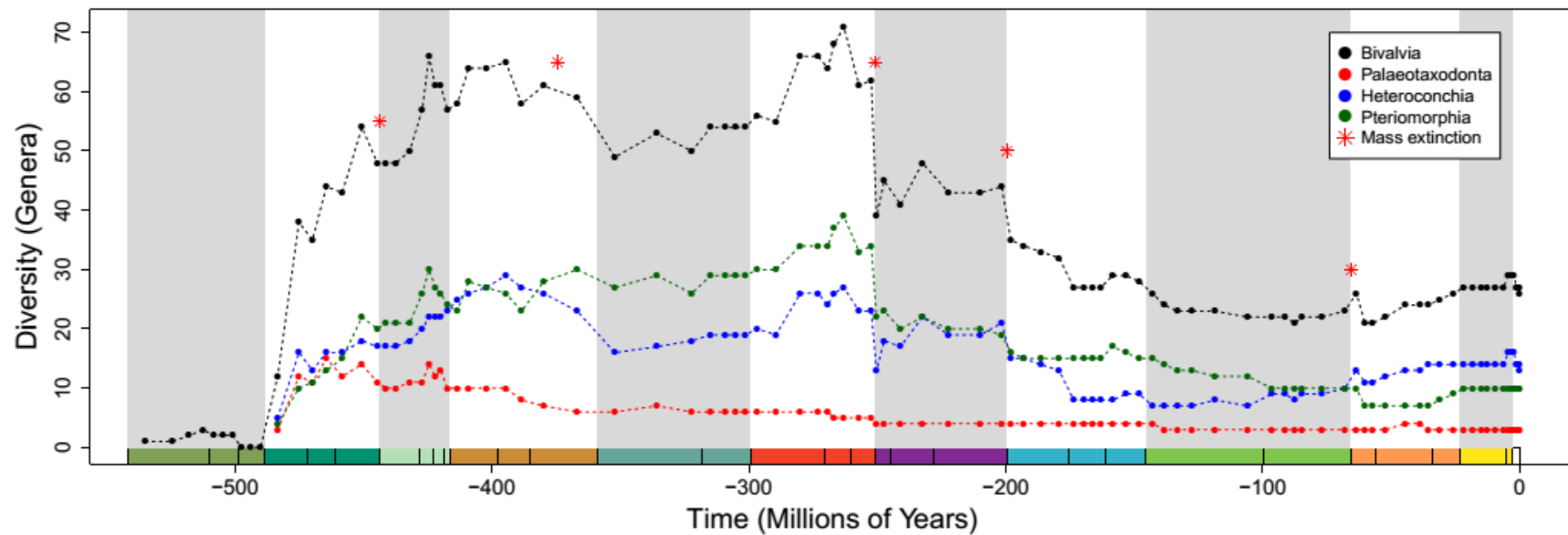


Figure.3.8. Diversity estimates (traditional counts) for Bivalvia, Protobranchia ( $\approx$  Palaeotaxodonta), Heteroconchia and Pteriomorpha across the Phanerozoic calculated from the genera of the Carter *et al.* Black: Bivalvia; red: Protobranchia ( $\approx$  Palaeotaxodonta); blue: Heteroconchia; and green: Pteriomorpha. Diversity estimates for each stage are centred on the midpoint of each stage. The top five Mass extinctions (End Ordovician, Late Devonian, End Permian, End Triassic and End Cretaceous) are highlighted with red stars. The stratigraphic ranges derive from the Mixed Range stratigraphic treatment.

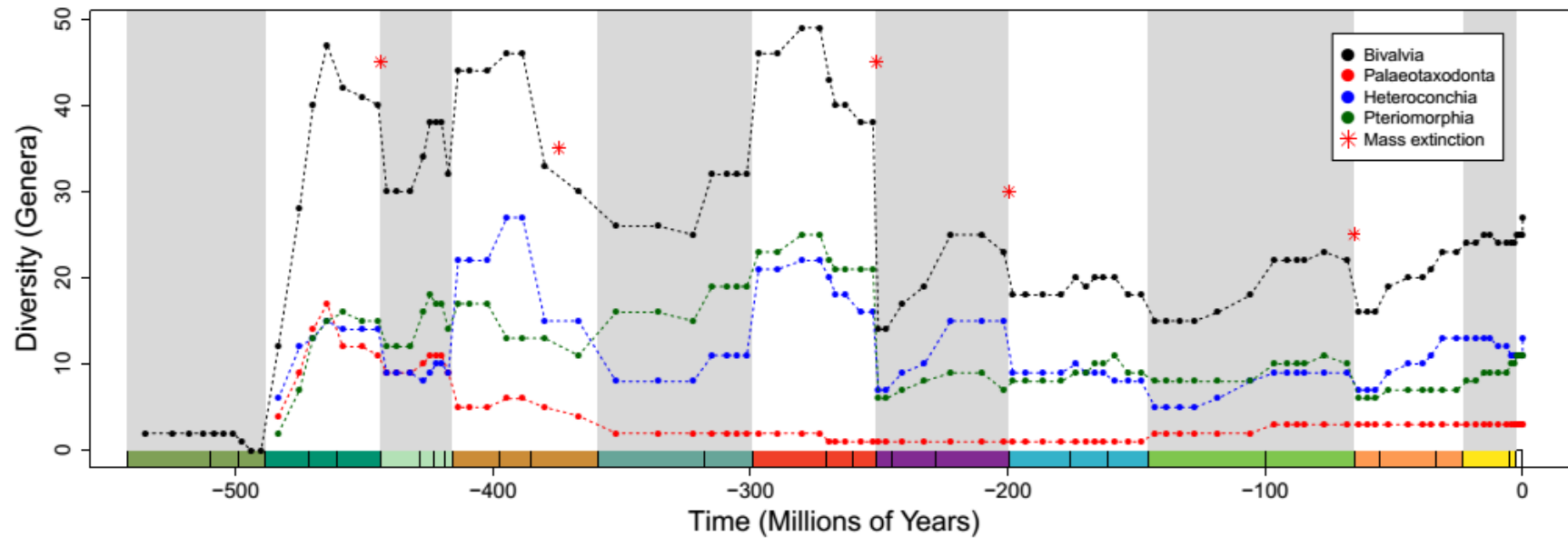


Figure.3.9. Diversity estimates (traditional counts) for Bivalvia, Protobranchia ( $\approx$  Palaeotaxodonta), Heteroconchia and Pteriomorpha across the Phanerozoic calculated from the genera of the Carter *et al* (2006). Black: Bivalvia; red: Protobranchia ( $\approx$  Palaeotaxodonta); blue: Heteroconchia; and green: Pteriomorpha. Diversity estimates for each stage are centred on the midpoint of each stage. The top five mass extinctions (End Ordovician, Late Devonian, End Permian, End Triassic and End Cretaceous) are highlighted with red stars. The stratigraphic ranges derive from the revised Treatise stratigraphic treatment.

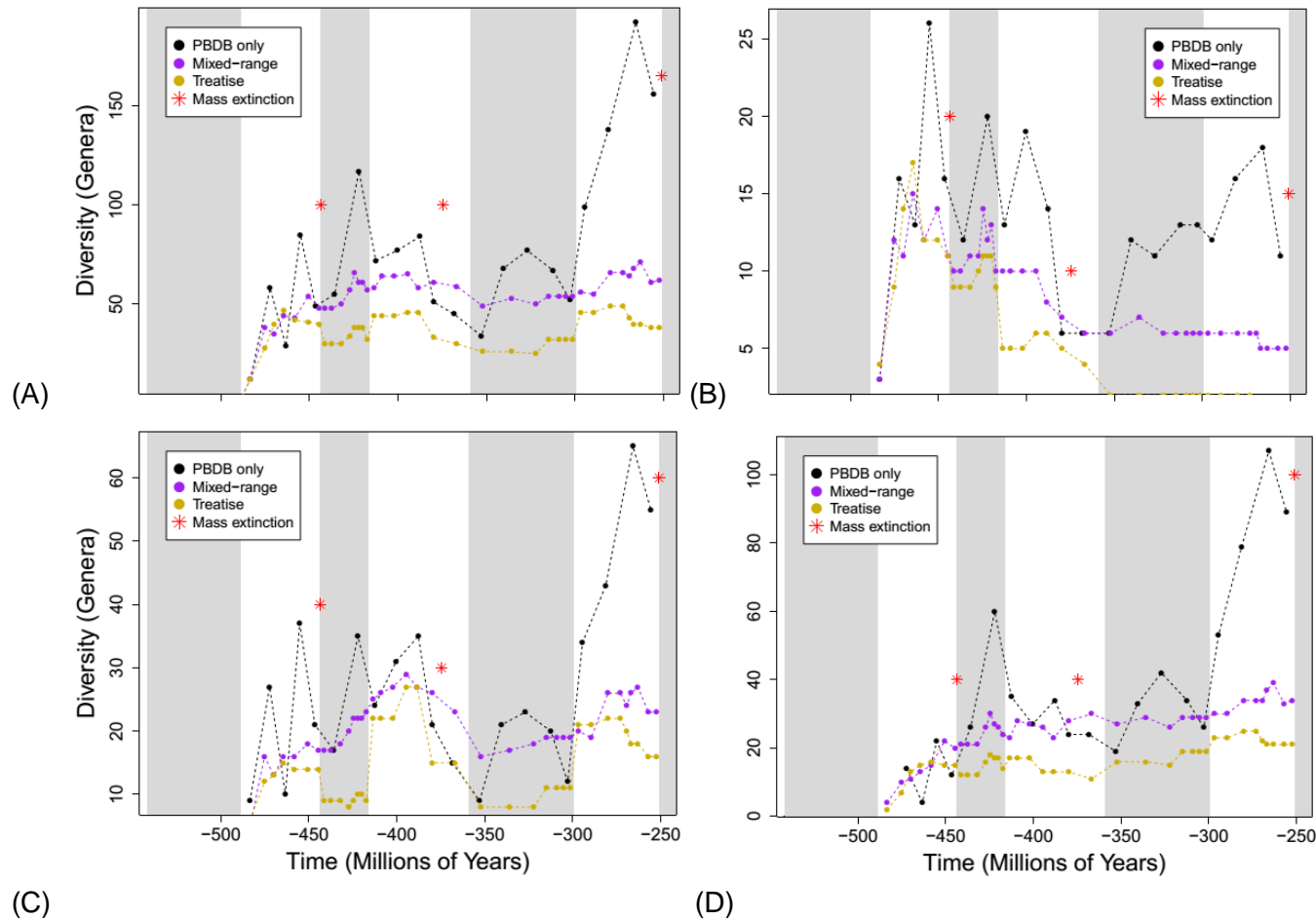


Figure.3.10: Comparing generic diversity (traditional counts) of bivalves across the Palaeozoic. (A) Bivalvia (B) Protobranchia ( $\approx$ Palaeotaxodonta). (C) Heteroconchia. (D) Pteriomorphia. Coloured lines represent the different sources of stratigraphic data. Black: counts from the Paleobiology Database (Kieissling et al. 2013); purple: counts of taxa in morphological data matrix using the Mixed Range treatment; yellow: counts of taxa in morphological data matrix using revised Treatise dataset. The time bins for the Paleobiology Database are split into intervals of roughly  $\sim 10$  million years. The time bins for taxa in the morphological matrix are the mid-points of each stage.

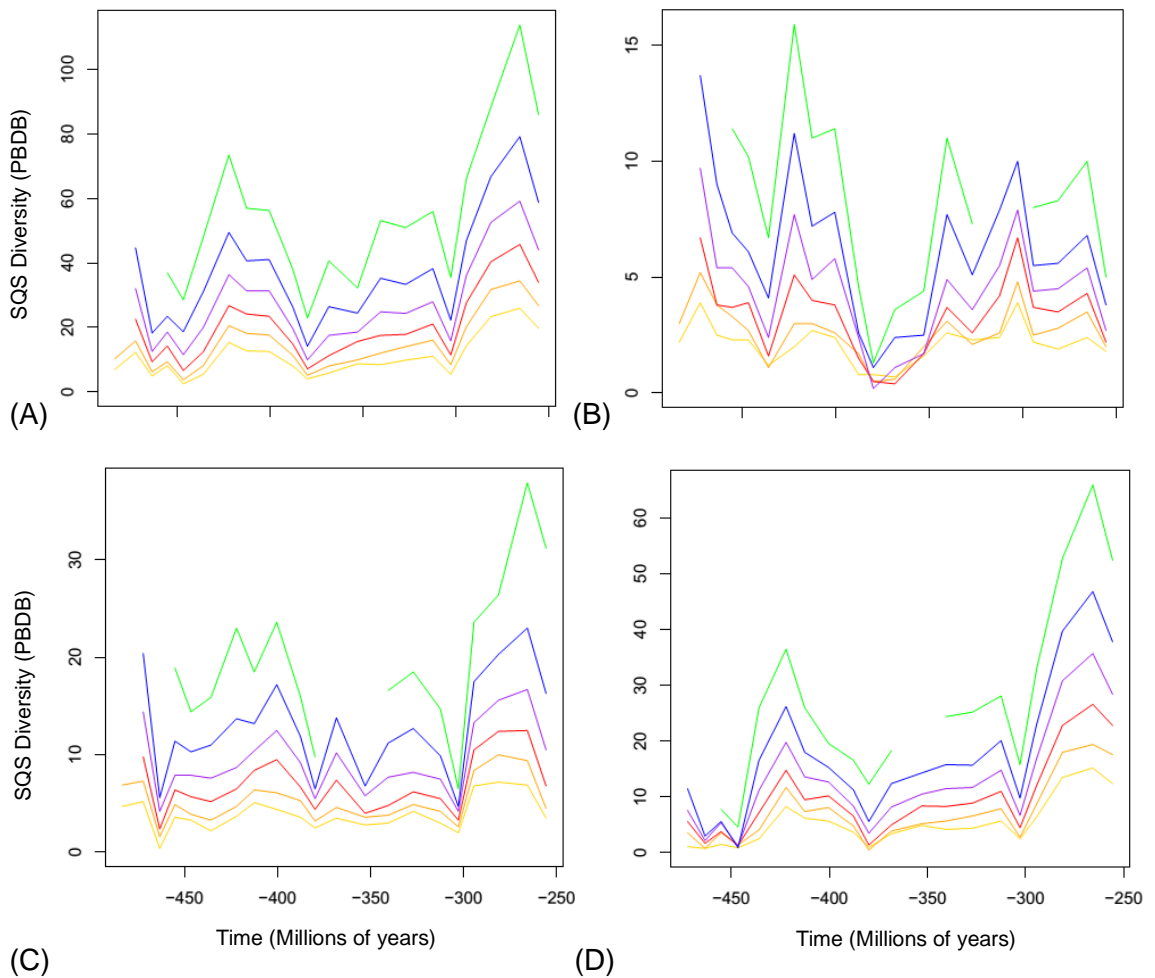


Figure.3.11. Diversity (SQS) of the Bivalvia and constituent subclades across the Palaeozoic using the Paleobiology Database online tools. (A) Bivalvia; (B) Protobranchia; (C) Heteroconchia; and (D) Pteriomorphia. Colours represent different “quorum” levels. Green: 0.9; blue: 0.8; purple: 0.7; red: 0.6; orange: 0.5; yellow 0.4.

from the three subclades having more genera in the Ordovician in the Treatise treatment than the Mixed Range Treatment.

Both treatments show a drop across the Devonian which continues into the Carboniferous and finally a rise in diversity in the Permian (Fig.3.8-9). The main differences between the treatments is in the first three periods, and stems from the Treatise containing more Heteroconchia genera (Fig.3.8-9 blue line) than Pteriomorphia (Fig.3.8-9 green line) in the Ordovician – Devonian compared to the Mixed Range treatment. The trajectories of subclades in the two treatments are similar in shape: Protobranchia reach their highest diversity in the Ordovician (Fig.3.8-9 red line) dropping across the rest of the Palaeozoic; the Heteroconchia follow a similar shape to the bivalve diversity curve (Fig.3.8-9 blue line) and the Pteriomorphia grow at a more rapid in the Ordovician or Silurian then more gradually until the end of the Permian (Fig.3.8-9 green line).



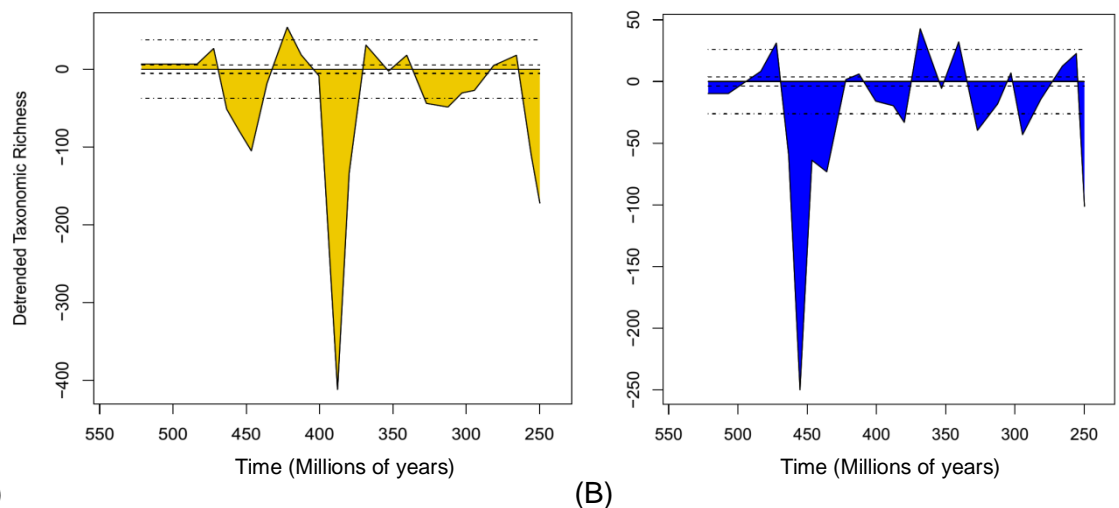


Figure.3.12. The detrended diversity of bivalves through the Palaeozoic after removing the effects of fossil record sampling biases using two sampling proxies. (A) diversity after removing the effect of the number of collections; and (B) the diversity after removing the effect of the number of formations. Outer dashed line represents the 95% confidence intervals of diversity purely the result of the sampling proxy (standard error). Inner dashed line reflects standard deviation. The gold and blue colourings are for aesthetic purposes.

The diversity (traditional counts) of Bivalvia and each subclade across the Palaeozoic for the Mixed Range and Treatise treatments were compared to the diversity based on the whole Paleobiology Database (Fig.3.10). This was done to assess the amount of missing diversity in the Palaeozoic specifically. In all treatments, the level of diversity was of the Mixed Range and Treatise treatments were smaller than the Paleobiology Database. The greatest difference is in the diversity in the Permian, where the Paleobiology Database contains many more taxa than found in data of Carter *et al.* (2000; 2006).

The shareholder quorum subsampling results for the Paleobiology Database can be seen in Figure 3.11. The data 0.9 quorum could not be reached for the individual parts of the curve but the 0.8 quorum level was reached in Bivalvia, Protobranchia, Heteroconchia and Pteriomorpha. The trajectory for Bivalvia (Fig.3.11A) shows a diversity spike in the Silurian, a diversity drop through the Devonian with a recovery and stability in the Carboniferous. There is a small extinction at the end of the Carboniferous and then diversity spike in the Permian prior to the extinction. This diversity increase was double the size of the one in the Silurian. The subclade SQS results show a similar pattern as the one for the Bivalvia as a whole.

The results of Figure 3.12 show the intervals of time where diversity is argued to be greater (positive values) or less than (negative values) expected from the fossil record sampling proxy (number of collections and formations respectively). The results show evidence for the Ordovician mass extinction (Fig.3.12B) and Devonian diversity drop (Fig.3.12A) as shown by large negative deviations from the expected diversity at these

times. There is evidence in the each graph for the other event, but the deviation is much smaller (but still outside the 95% limits of diversity expected by fossil record bias alone). Both sets of results (Fig.3.12A and B) find some deviation from the 95% confidence limits in the Carboniferous, but at slightly different times. Both find the drop in diversity with the end Permian extinction. Positive deviations outside of the expected zone were small, occurring in the Silurian for the number of collections (Fig.3.12A) and at the beginning of the Ordovician, End Devonian and Early Carboniferous (Fig.3.12B).

### **3.17 Morphospace Occupation Across the Palaeozoic**

The morphospace plots of Figure.3.13 show the occupation of all bivalve genera originating in the Palaeozoic. There are three distinct clusters corresponding to the three major clades, but not all regions are occupied at any one time. Each morphospace is built from one of two pools of stratigraphic ranges: the Mixed Range treatment (Paleobiology Database, Krug *et al.* and other literature) and the Treatise treatment (Treatise of Invertebrate Palaeontology). These two pools differ in the assignment of stratigraphic ranges for many genera. This is due to the more conservative nature of the Treatise of Invertebrate Palaeontology compared to the Paleobiology Database in assigning first and last dates. The manner in which the three subclades evolve through the morphospace remains the same with either stratigraphic treatment. It is worth noting the subtle effects these stratigraphic differences have on the occupation of morphospace, morphospace clustering and disparity.

The most significant difference between the two pools is the number of taxa assigned to a Palaeozoic origination. The Mixed Range treatment suggests 198 genera originated by the end Permian mass extinction, whilst 185 genera are record in the Treatise. The thirteen genera found in the Mixed Range treatment that are missing from the Treatise dataset are: the genus *Nucula* (Protobranchia); the genera *Astarte*, *Cardinia*, *Myophoria*, and *Pleuromya* (Heteroconchia); and the genera *Mytilus*, *Rhombopteria*, *Entolioides*, *Entolium*, *Syncyclonema*, *Pleuronectites*, *Oxytoma* and *Pinna* (Pteriomorpha).

The Treatise treatment provides more conservative range estimations on several genera compared to the Mixed Range treatment (primarily the product of the Paleobiology Database). The some of the thirteen genera missing from the Treatise treatment are cosmopolitan, with far reaching evolutionary histories. As fossils are not perfectly preserved in every instance, it is inevitable that some specimens would be misidentified as part of these taxa, extending their stratigraphic range. The Treatise has made every attempt to minimise miss-classifying specimens, and identifying/correcting synonyms. The two pools can be considered the conservative (Treatise) and liberal (Mixed Range) approach to the same data. The differences in the morphospace distributions of genera through time in the two treatments were mostly insignificant, with each space mirroring the

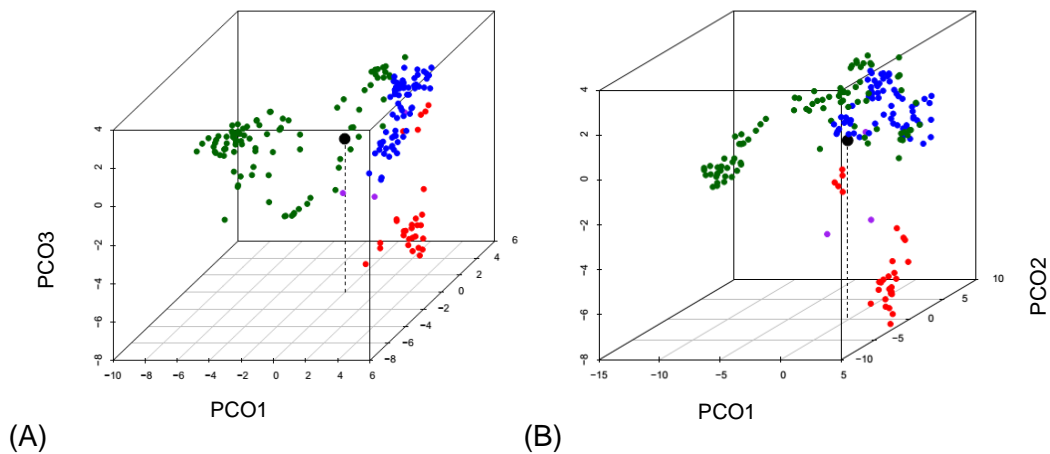


Figure.3.13: Occupation of bivalves in morphospace along the first three principal coordinates. The morphospace is occupied by all taxa originating in the Palaeozoic according to: (A) the Paleobiology Database (PBDB), and (B) the Treatise of Invertebrate Palaeontology. Colours represent subclades whilst all points represent Bivalvia across the Palaeozoic: Protobranchia: red; Heteroconchia: blue; and Pteriomorphia: green. The black dot represents the global centroid. Dotted lines are a guide to the distribution of clades along principal coordinate axis one and two.

other. When significant differences between the treatments did arise, the reasons behind these discrepancies have been explored.

From their origins in the Cambrian, bivalves orbit the global centroid (centroid of the entire Palaeozoic bivalve morphospace). The area of space occupied by the Cambrian bivalve *Fordilla* is near a region that contains genera from all three major clades at different times throughout the Palaeozoic. The three clades never overlap at any given time. There is therefore close convergence of some genera within all three clades on similar body plans. Protobranchia occupy two sub-regions of space throughout their history, showing most variation along the third PCO axis. They also explore the edges of their space quickly (Early Ordovician).

The timing of this filling is subtly different for the Mixed Range (Tremadocian) and Treatise datasets (Floian). For the entire Palaeozoic new protobranchian genera originate in these two subregions but not between. Heteroconchia shows most variation along PCO 2. Unlike the Protobranchia, they expand out from their initial starting position (through the Ordovician), but then contract inwards and then expand again until the end Devonian. They then change little for the rest of the Palaeozoic.

The Pteriomorphia show the greatest expansion through the space. The only difference between the two treatments is the direction of expansion (due to the mirroring effect), not the pattern of expansion. The following account is for the Mixed Range treatment. Pteriomorphia initially occupy space near the Heteroconchia, exploring PCO 2 (like the Heteroconchia) but progressively move towards the negative values of PCO 1. From their starting position (which they occupy for the entire Palaeozoic), they spread out

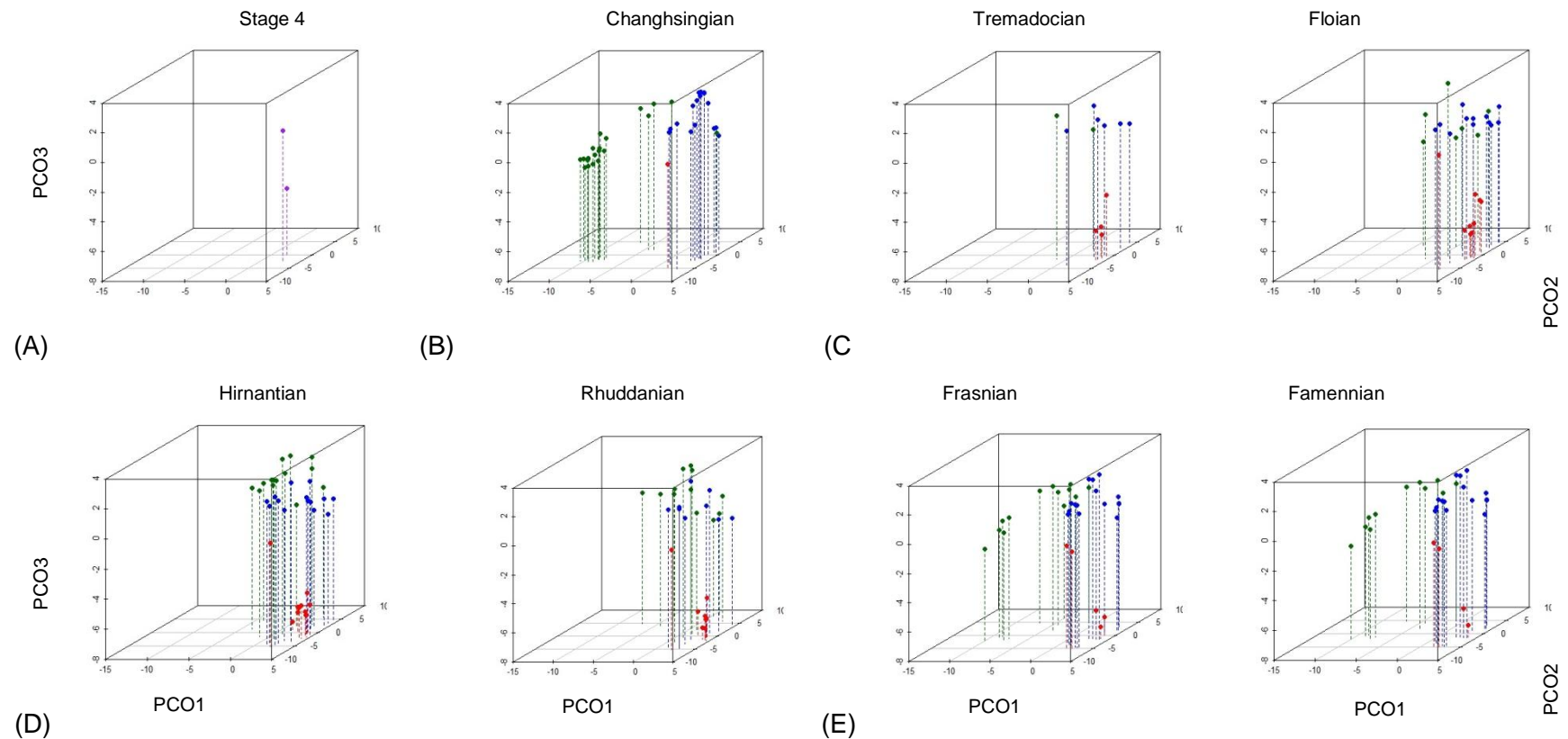


Figure.3.14. Morphospace occupation of bivalves along the first three principal coordinates across selected stages of the Palaeozoic, showing the distribution of taxa across specific time intervals (Treatise treatment). (A) Cambrian: Stage 4 (~515-510 million years ago). (B) End Permian: Changhsingian boundary (253.8-251million years ago). (C) Early Ordovician: Tremadocian-Floian (488.3-471.8 million years ago). (D) End Ordovician: Hirnantian-Rhuddanian boundary (445.6-439million years ago). (E) Late Devonian: Frasnian-Famennian boundary (385.3-359.2million years ago). Purple: Cambrian genera; red: Protobranchia; blue: Heteroconchia, and green Pteriomorphia. Dotted lines are a guide to the distribution of clades along principal coordinate axis one and two. All Stages found in Appendix.2.

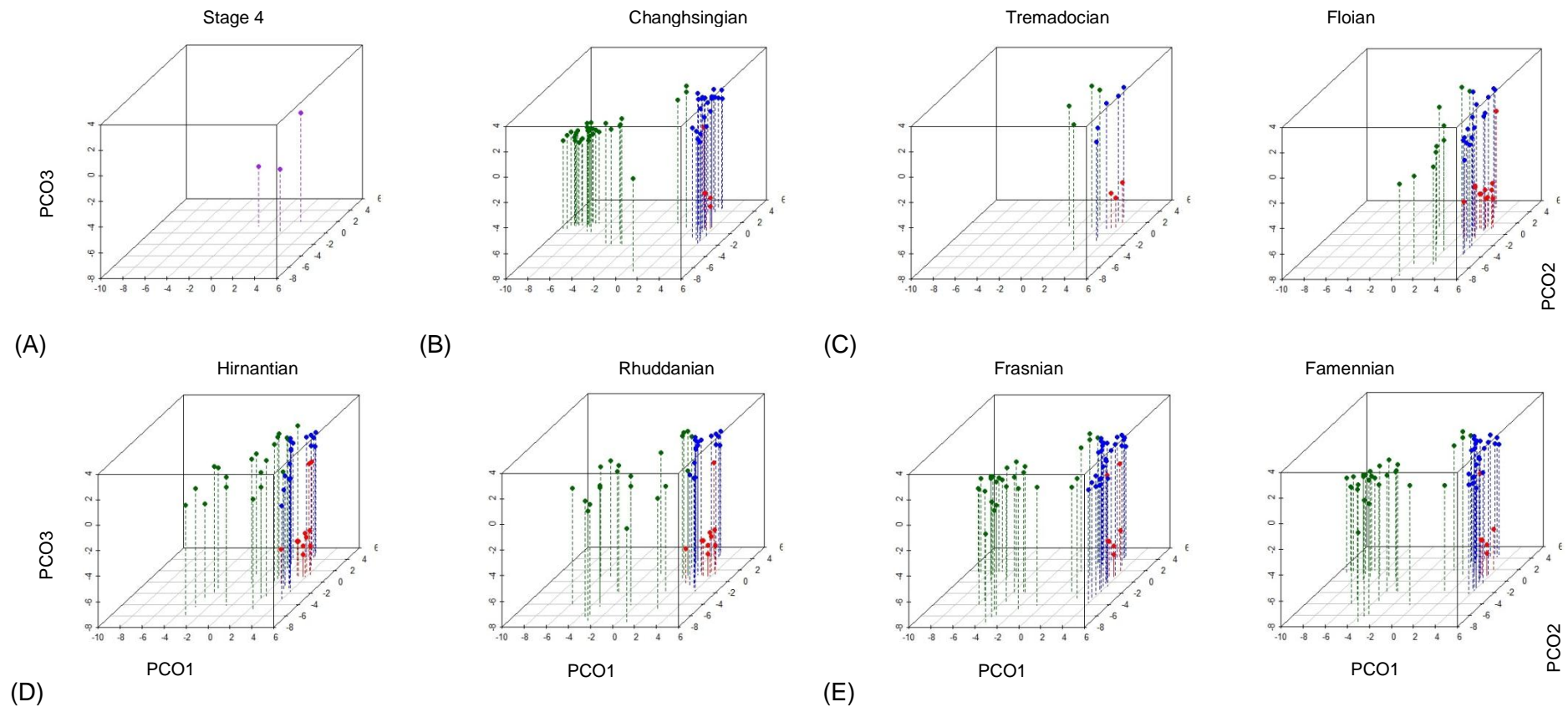


Figure.3.15. Morphospace occupation of bivalves along the first three principal coordinates across selected stages of the Palaeozoic, showing the distribution of taxa across specific time intervals (Mixed Range treatment). (A) Cambrian: Stage 4 (~515-510 million years ago). (B) End Permian: Changhsingian boundary (253.8-251million years ago). (C) Early Ordovician: Tremadocian-Floian (488.3-471.8 million years ago). (C) End Ordovician: Hirnantian-Rhuddanian boundary (445.6-439million years ago). (D) Late Devonian: Frasnian-Famennian boundary (385.3-359.2million years ago). Purple: Cambrian genera; red: Protobranchia; blue: Heteroconchia, and green Pteriomorpha. Dotted lines are a guide to the distribution of clades along principal coordinate axis one and two. All Stages found in Appendix.2.

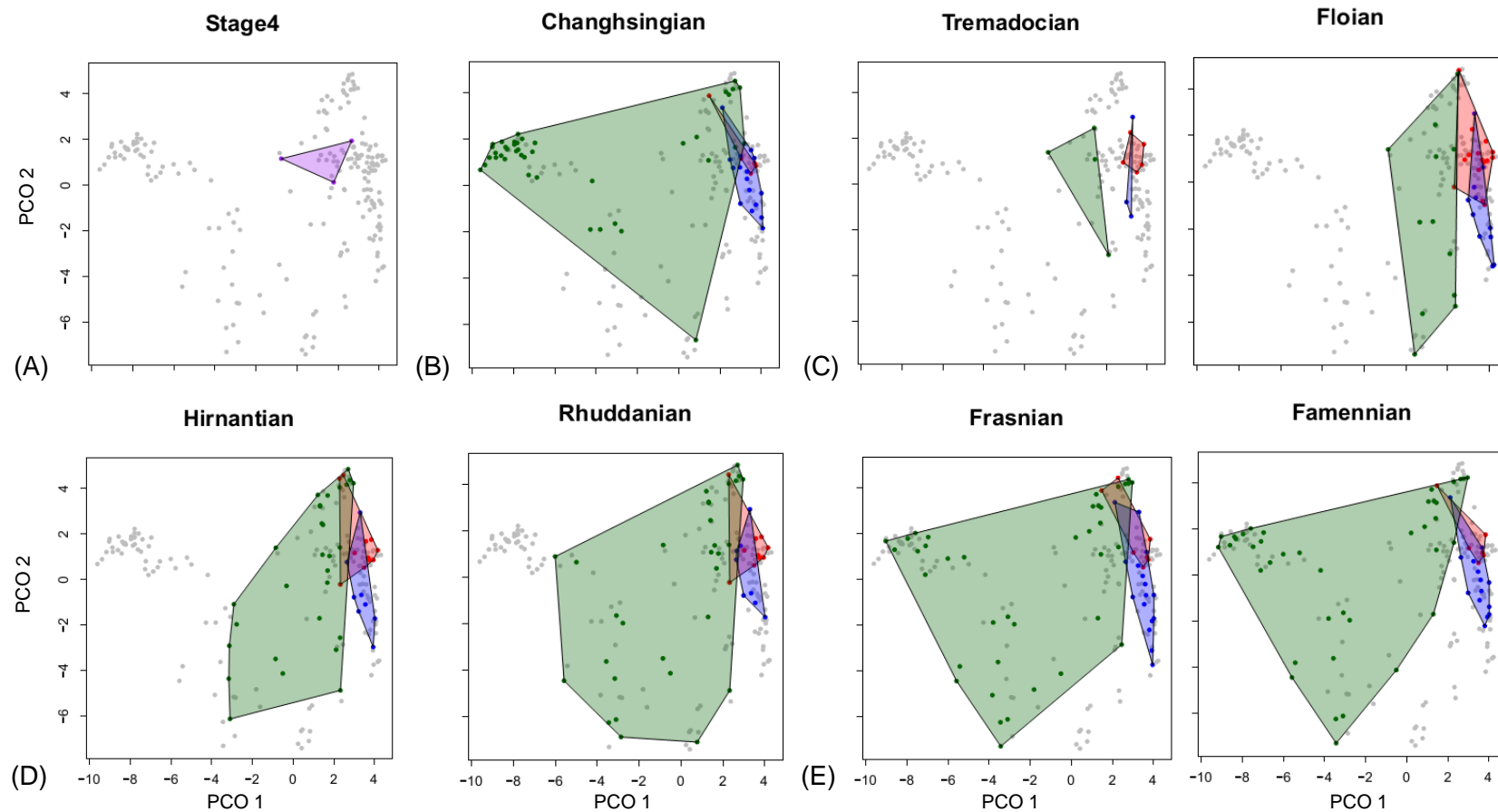


Figure 3.16: Morphospace occupation (convex hull) of bivalve genera across the Palaeozoic (Mixed Range). Morphospace produced using all characters. (A) Cambrian Stage 4 (~515-510 million years ago). (B) Changhsingian boundary (253.8-251million years ago). (C) Tremadocian-Floian: Early Ordovician (488.3-471.8 million years ago). (D) Hirnantian-Rhuddanian boundary (445.6-439million years ago). (E) Frasnian-Famennian boundary (385.3-359.2million years ago). Grey dots represent total morphospace occupation. Purple: Cambrian bivalve genera; red: Protobranchia; blue: Heteroconchia, and green Pteriomorphia. The x and y-axes are the first and second principal coordinates. (Appendix.2)

in morphospace. The majority of genera occupy the extremes of the morphospace cloud that expands through the Palaeozoic, leaving only a modest number of genera in a sparsely populated region its wake. These remaining genera also go extinct, leaving gaps between the more populated areas of the space. As the Pteriomorphia expand away from their point of origin, old areas thin out. This results in the creation of two distinct clusters after the Hirnantian (Late Ordovician).

In the Mixed Range treatment, one taxon (*Pterinea*) is left behind (Early Carboniferous) as the cloud furthest from the point of origin moves toward negative values of PCO1 / positive values of PCO2 and 3. This taxon remains on its own until it goes extinct near the end of the Carboniferous. In the late Early Permian an adjacent area of the space close to where *Pterinea* previously vacated becomes re-occupied by *Cosmetodon* (Artinskian – Late Early Permian). The Treatise finds the same replacement, except it occurs earlier with a longer gap between the two genera: *Pterinea* goes extinct in the Early Devonian (Emisian) and *Cosmetodon* originates in the Middle Carboniferous (Bashkirian). Generally the Ordovician shows an increase in the amount of morphospace occupied by all three major clades. Pteriomorphians explore the most, despite remaining a relatively diffuse but singular cluster until the end of the Ordovician (Fig.3.14-15D).

Heteroconchs explore mainly within the confines of the space they establish by the first two stages of the Ordovician (Tremadocian and Floian (Fig.3.14-15C, Fig.3.16C)). The Silurian shows an increase in diversity across all three clades whatever method is used, and this is reflected in the space. This does not result in any fundamentally new portions of the space becoming occupied; all exploration is within the areas currently colonised. The Devonian sees two prominent clusters emerge for the Pteriomorphia; one at their origin, the other some distance from it. Heteroconchs remain near their centre of origin. After a contraction in the late Ordovician, they briefly reoccupy this space in the mid Devonian but then contract again, remaining near their origin for the rest of the Palaeozoic.

### **3.18 Morphospace Occupation Across the Phanerozoic**

Serious selection bias in favour of Palaeozoic originating genera makes interpretation of post Palaeozoic morphospace occupation and disparity difficult. It is only considered here in a preliminary fashion, with the caveat that the addition of taxa originating in the Mesozoic/Cenozoic could change the results very significantly. After the End Permian extinction, the morphospace (Mixed Range and Treatise treatments) becomes sparser in the centre of the space and more clustered for each subclade (Fig.3.16AD). The Protobranchia remain in their own cluster near the origin point of the group, as do subsections of the Heteroconchia and Pteriomorphia. By the Middle Jurassic (Aalenian), the Heteroconchia have formed two clusters at opposite ends of the

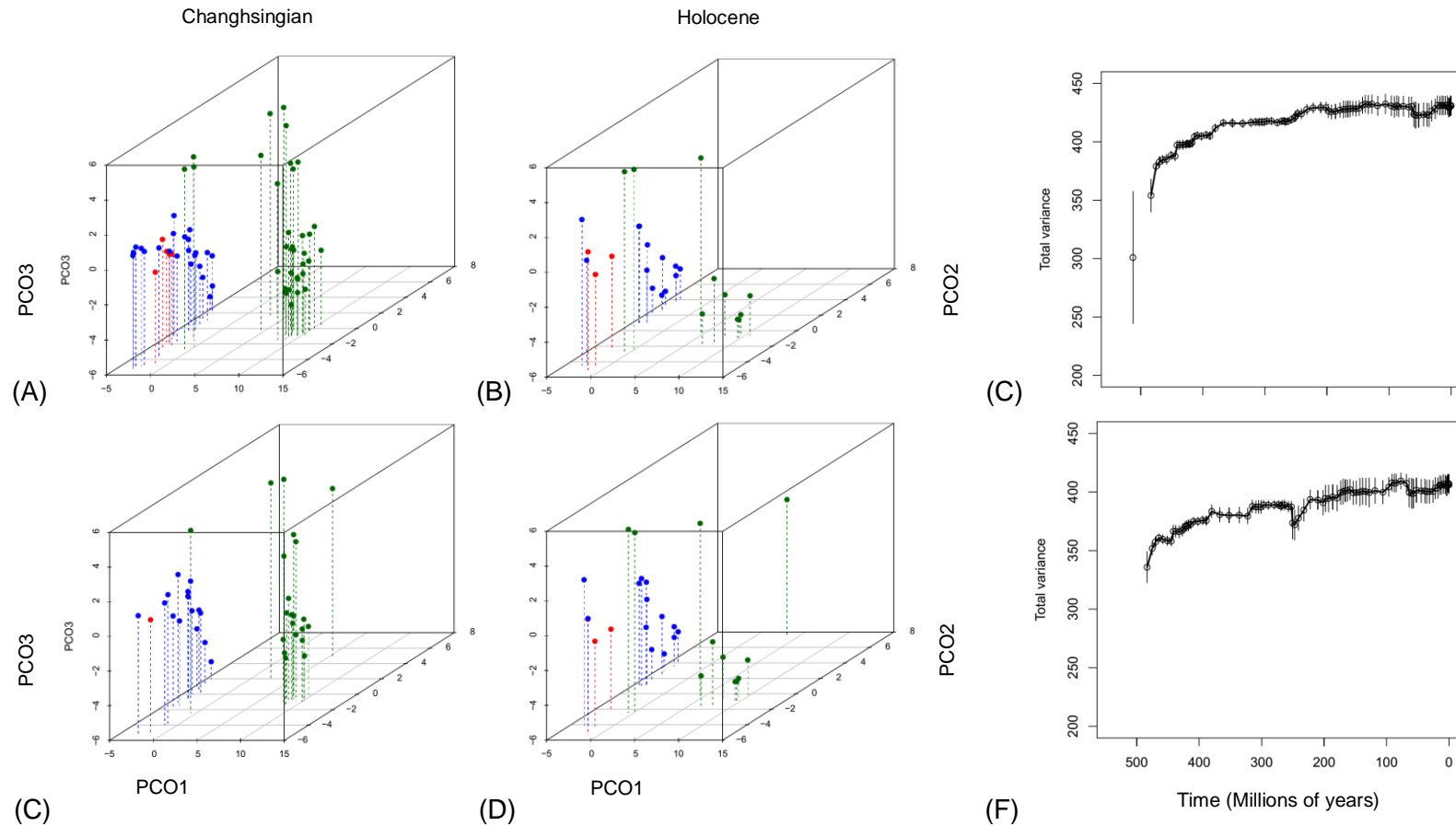


Figure 3.17: Morphospace occupation and disparity in post-Palaeozoic bivalves. (A) Morphospace occupation at the End Permian boundary (Changhsingian) (Mixed Range). (B) Morphospace occupation of extant genera (Mixed Range). (C) Disparity (sum of variances) of Phanerozoic bivalves (Mixed Range). (D) Morphospace occupation at the End Permian boundary (Treatise). (E) Morphospace occupation of extant genera (Treatise). (F) Disparity (Treatise). Red: Protobranchia genera; blue: Heteroconchia, and green Pteriomorpha. Dotted lines are a guide to the distribution of clades along principal coordinate axis one and two. Black line: mean sum of variance, black error bars: standard error.



morphospace they occupied at the End of the Permian. The intermediate space remains unoccupied up to the Recent (Fig.3.17BE).

The Pteriomorphia remain in two clusters defined at the end of the Permian with taxa originating/going extinct within these. The Treatise dataset differs from the Mixed Range treatment in that the Heteroconchia closest to origin point of the clade forms a broader cluster (Fig.3.17BE) From the End Permian onwards; no new areas of morphospace are explored up to the Recent. Total disparity continues to increase in steady but shallow increments up to the Recent (Fig.3.17CF). Disparity in individual subclades does not change over this time. Diversity estimates vary between pre- and post-Palaeozoic bivalves, (taxa in the morphological matrix) the latter decreasing from the End Permian up to the Recent.

### **3.19 Effects of Character Selection on Morphospace Occupation & Disparity**

A number of characters pertain to morphological variation in particular subclades or even particular orders. For example, characters 52, 53 and 54 refer to shell shapes within Pteriomorphia, Protobranchia and Solemyida respectively. In this case the other genera are coded with 0's. If many characters were chosen to pull apart the phylogenetic detail within some clades over others, this could artificially increase the relative disparity of those clades, and pull the global centroid toward those taxa split by more characters. The effect that character selection may have on the pattern of disparity and occupation of genera in morphospace was therefore explored.

The first set of treatments considered including removing characters specific to particular subclades: (1) characters varying across Protobranchia only (253-215=38 removed), (2) characters varying across Heteroconchia only (253-230=23 removed), (3) characters varying across Pteriomorphia only (253-163=90 removed). Characters specific to particular subclades were identified by isolating characters that were variable only in one subclade and invariable the remaining two subclades. The next treatment removed all characters that were invariable for any one subclade, leaving only those characters that varied across all subclades (253-96=157 removed). The final treatment removed those characters considered to have *a-priori* assumptions regarding character evolution (253-106=147 removed). These characters entailed character descriptions and states referring to specific key words namely 'primitive', 'ancestral', 'advanced', and 'derived'.

The results of the different treatments can be seen in Figure 3.18A-E. The graphs show the subclade distribution in morphospace using PCO axes 1-3, the morphospace occupation (convex hull) of each subclade using PCO axes 1 and 2 and the trend in disparity through time. Across all the treatments, the shape of the entire morphospace does not change fundamentally. The subclades still form three clusters, of which Pteriomorphia explores the greatest amount. There is a rotation of the morphospace in each treatment, but this is only an artefact of the ordination process. The substantial

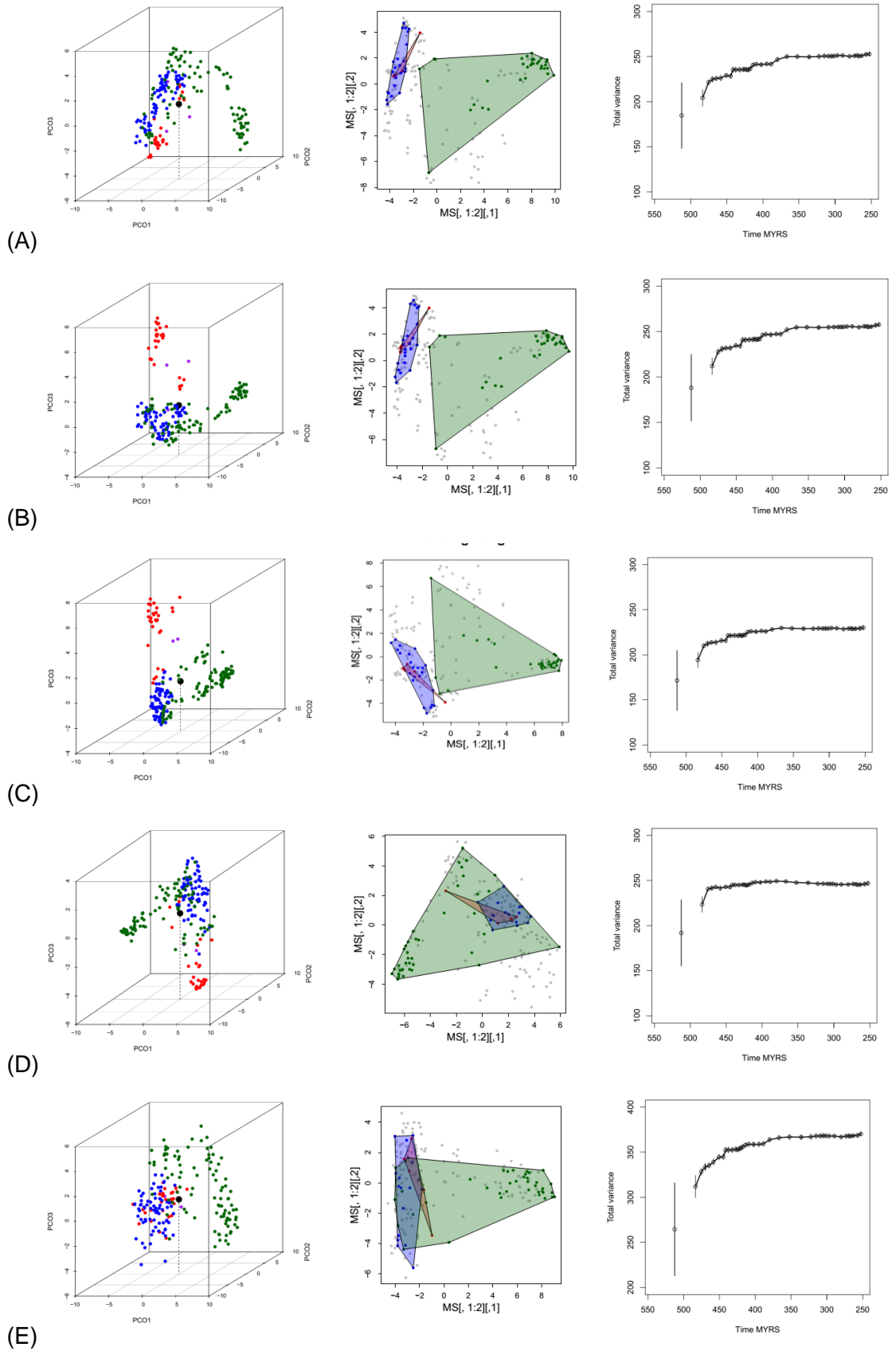


Figure.3.18: The effect of character choice. Left panel: total morphospace for the Palaeozoic, PCO axes 1-3. Red: Protobranchia; blue: Heteroconchia; green: Pteriomorphia. Middle panel: morphospace occupation (convex hull) using PCO axes 1-2 at the End Permian Changhsingian boundary. Colours same as in (A). Right panel: disparity (sum of variances) across the Palaeozoic. Black line: mean; error bars: standard error. (A) Protobranchia specific characters removed. (B) Heteroconchia specific characters removed. (C) Pteriomorphia specific characters removed. (D) All invariable characters removed. (E) *A-priori* character assumption removed.

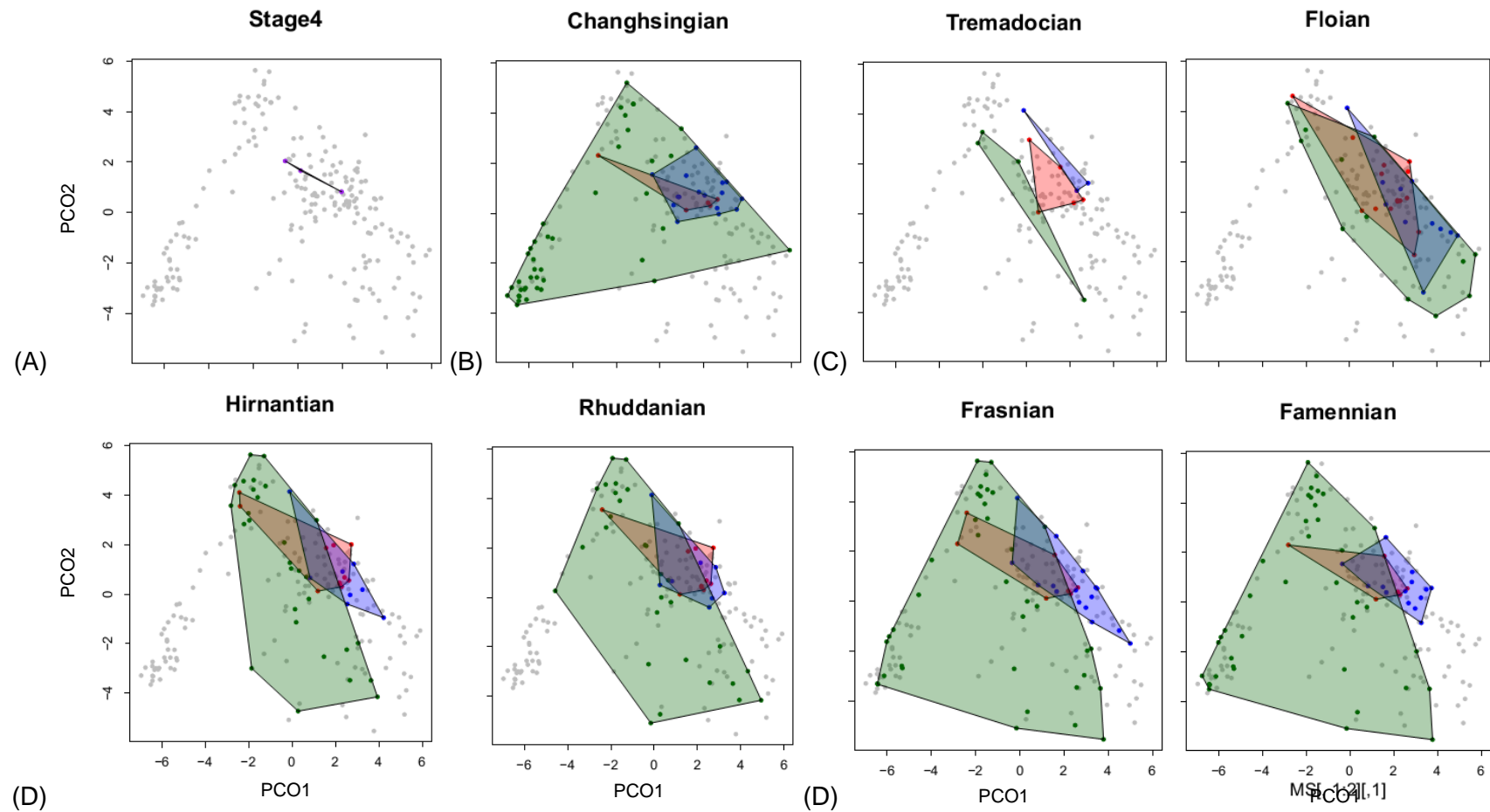


Figure.3.19: Morphospace occupation (convex hull) of Bivalvia across the selected parts of the Palaeozoic (Mixed Range). Morphospace produced using the 96 character matrix (all invariant characters removed). (A) Cambrian Stage 4 (~515-510 million years ago). (B) Changhsingian boundary (253.8-251million years ago). (C) Tremadocian-Floian: Early Ordovician (488.3-471.8 million years ago). (D) Hirnantian-Rhuddanian boundary (445.6-439million years ago). (E) Frasnian-Famennian boundary (385.3-359.2million years ago). Grey dots represent total morphospace occupation. Purple: Cambrian bivalve genera; red: Protobranchia; blue: Heteroconchia, and green Pteriomorpha. The x and y-axes are the first and second principal coordinates.

difference between the character treatments is the degree to which subclades overlap. As more characters related to specific subclades are removed, the greater the overlap in morphospace between subclades as a whole (Fig.3.18) and through time (Fig.3.19). The treatments do not affect the pattern of disparity through time (Fig.3.18).

### **3.20 Clustering Intensity Across the Phanerozoic**

#### **3.21 Nearest Neighbour Analysis**

The results of the nearest neighbour analysis for are shown in Fig.3.20, 3.21 and 3.22. The intensity of clustering across bivalves as a whole is shown in Fig.3.20A & C. These two graphs are built upon the Mixed Range and Treatise treatments respectively. Both show a high level of clustering throughout the Palaeozoic. There are different patterns in clustering depending on the treatment used. Both treatments show an initial jump in clustering intensity in the Early to Middle Ordovician. Whereas the Mixed Range treatment shows a gradual increase in clustering intensity across the Palaeozoic, the Treatise treatment shows a more complex pattern.

The latter treatment suggests that after the Early-Middle Ordovician high, there is a drop through the Late Ordovician. In the Early Silurian (Llandovery epoch) clustering intensity remains stable, and then increases across the Middle Silurian boundary. The intensity drops once more in the Late Silurian to the level in the Early Silurian. In the Early Devonian, clustering intensity increases again, remaining stable until the Middle-Late Devonian. Cluster intensity then drops in the Middle-Late Devonian/Early Carboniferous, remaining stable for the rest of the Carboniferous period. In the Early Permian, the clustering intensity increases again, remaining stable up to the end of the Palaeozoic.

The rapid increase in the Ordovician corresponds well with the establishment of the three subclades in the Early Ordovician (Fig. 24). Each subclade shows its own unique pattern of clustering. Figure 3.20B&D show the clustering of the subclades built from the total bivalve morphospace (20 axes) using the Mixed Range and Treatise treatments. Figures 3.21 and 3.22 also show subclade clustering intensity, but the morphospace of each is built in isolation. The difference between Figs.3.21 and 3.22 is the number of axes used to calculate clustering. Figure 3.21 shows the clustering intensity based on 10, 15 and 20 PCO axes for the Protobranchia, Heteroconchia and Pteriomorphia respectively. These values were chosen to reduce the effect of the non linear relationship between raw distances (character dissimilarity) and subsequent addition of PCO axes. Figure 3.22 shows the same data using the maximum number of axes that can be calculated using uncorrected PCO. The results described here are for Fig.3.21, but the same relationships are seen in Figure 3.22.

The differences between Fig.3.20 and Fig.3.21 are mainly an artefact of the scope of each treatment. The pattern of clustering in each subclade reflects the position of genera within those subclades. When clustering within each subclade is regarded

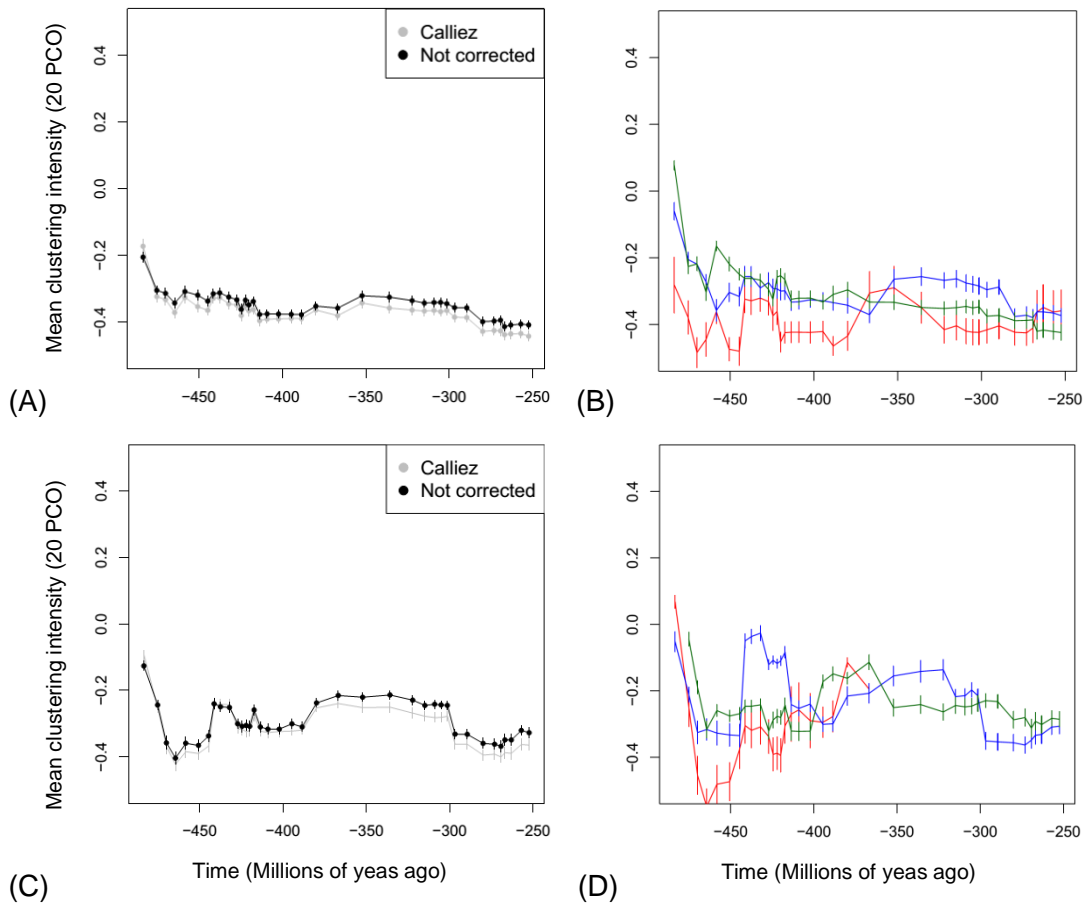


Figure.3.20. Intensity of clustering of subclades across the Palaeozoic based on nearest neighbour analysis of the first 20 PCO axes. Red: Protobranchia, blue: Heteroconchia, and green: Pteriomorpha. (A) Intensity of clustering based on morphospace built using the Mixed Range treatment. (B) Intensity of clustering of the three subclades based on the morphospace of Bivalvia (Mixed Range). (C) Intensity of clustering based on the morphospace built using the Treatise dataset. (D) Intensity of clustering of the three subclades based on the morphospace of Bivalvia (Treatise). Error bars are the standard error based on 20 replications

together, the relative values of each subclade become more negative. This happens because each subclade belongs to its own, non-overlapping area of the total morphospace. When considered together (Fig.3.20B&D), the clustering of each subclades intensifies. This is analogous to looking at the map of a town, showing the position of buildings. In isolation the town shops may seem spread out and randomly placed, but when the surrounding area and other towns are simultaneously considered, the shops in the original town will become relatively more clustered.

The Protobranchia becomes most intensely clustered in the early Ordovician (Fig.3.21A). Across the End Ordovician boundary (443.7 million year ago), there is a drop in clustering intensity. Clustering intensity increases gradually across the Silurian, but begins to drop across the Devonian, reaching its lowest level in the late Devonian - Early Carboniferous. Clustering begins to increase across the Early Carboniferous, and then

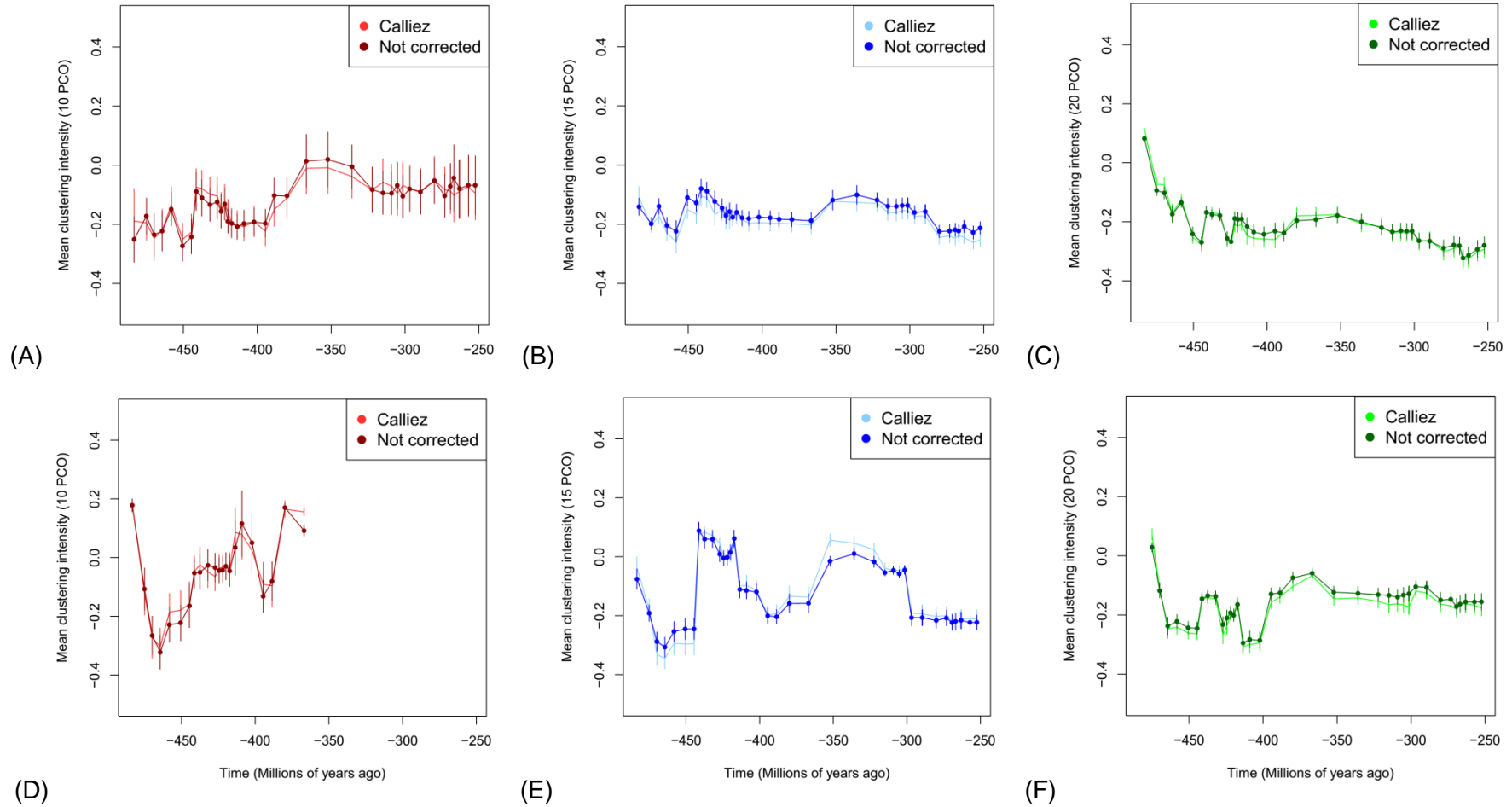


Figure.3.21. Clustering Intensity across the Palaeozoic using nearest neighbour analysis. The numbers of PCO axes are: 10, 15 and 20 axes for Protobranchia, Heteroconchia and Pteriomorphia respectively. Cailliez corrected PCO results highlighted in lighter colour than uncorrected PCO axes. (A) Clustering intensity of Protobranchia using the Mixed Range treatment. (B) Clustering intensity of Heteroconchia using the Mixed Range treatment. (C) Clustering intensity of Pteriomorphia using the Mixed Range treatment. (D) Clustering intensity of Protobranchia using the Treatise treatment. (E) Clustering intensity of Heteroconchia using the Treatise treatment. (F) Clustering intensity of Pteriomorphia using the Treatise treatment.

fluctuates with shorter periods of increase and decrease for the rest of the Palaeozoic. During this time Protobranchia never reach the same level of clustering intensity as in the Ordovician or Silurian. Overall, the Protobranchia have a two step clustering phase, maximum clustering attained early and reduced clustering thereafter. The major shifts in clustering coincide with the mass extinction events and increases in diversity after them.

Clustering intensity in the Heteroconchia (Fig.3.20B&D & 3.21B) is different to the pattern seen in Protobranchia. The latter is split by a clustered (but variable) and less clustered state corresponding to the first and second half of the Palaeozoic. The Heteroconchia, after an initial burst of clustering in the early Ordovician, remain in a stable clustered state except for two periods where clustering drops and stabilises. The first drop is across the End Ordovician - Early Silurian boundary, where clustering remains lower for the duration of the Silurian. After this period, clustering intensity reaches Ordovician levels across the Devonian, dropping again across the End Devonian - Early Carboniferous, and remaining at the new level for the remainder of that Period. The final switch back to Ordovician levels occurs across the Late Carboniferous/Early Permian boundary, where it remains at the same level for the rest of the Palaeozoic.

The clustering intensity curve for Pteriomorphia (Fig.3.20B&D & 3.21C) is the most similar to the curve for bivalves as a whole (Fig.3.20A&C) using both the Mixed Range and Treatise treatments. This is partly the result of their greater abundance compared to the other two clades, and because the Pteriomorphia explore a greater portion of the total morphospace compared to the other two subclades. In both Fig.3.20BD and 3.21C, Pteriomorphia are clustered to some degree, with a gradual increase in clustering over time. This increase in clustering reflects exploration of Pteriomorphia through the space (i.e., new taxa at the edge furthest from the origin and old taxa in intermediate areas of the space becoming extinct). There are only small differences between the intensity of clustering based on the Mixed Range and the Treatise datasets. The clustering for all bivalves (Fig.3.20C) is most intense in the Early Ordovician. The small fluctuations of clustering intensity in Fig.3.20A are seen in Fig.3.20C but are more exaggerated. This results in sharp drops in clustering intensity at the End Ordovician and Late Devonian. After the End Ordovician the clustering drops but begins to increase again until the Late Devonian. The Carboniferous shows less intense clustering and greater stability until the Early Permian where clustering intensifies once more, remaining at the same level until the End Permian.

There are some differences between the clustering intensity of subclades based on the Mixed Range (Fig.3.20B) and Treatise treatments (Fig.3.20D). The two patterns for Protobranchia have the most obvious differences. In Fig.3.20D/3.21D clustering intensity cannot be computed due to the lack of diversity according to the Treatise treatment. The clustering intensity for Protobranchia in the early half of the history is otherwise similar.

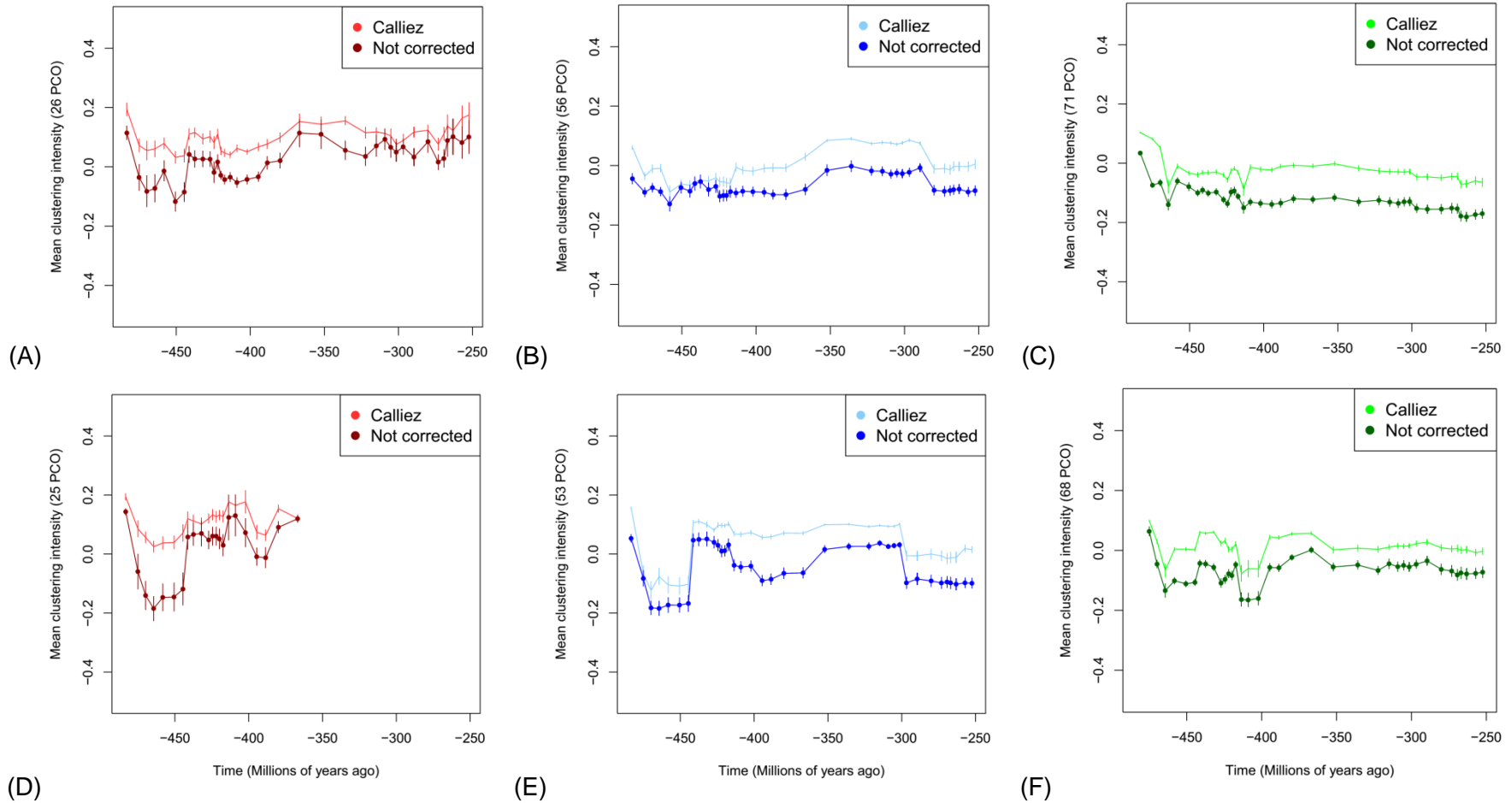


Figure.3.22. Clustering Intensity across the Palaeozoic using nearest neighbour analysis. The number of axes PCO for each subclade reflects the number of axes calculable using uncorrected PCO. Cailliez corrected PCO results highlighted in lighter colour than uncorrected PCO axes. (A) Clustering intensity of Protobranchia using the Mixed Range treatment. (B) Clustering intensity of Heteroconchia using the Mixed Range treatment. (C) Clustering intensity of Pteriomorpha using the Mixed Range treatment. (D) Clustering intensity of Protobranchia using the Treatise treatment. (E) Clustering intensity of Heteroconchia using the Treatise treatment. (F) Clustering intensity of Pteriomorpha using the Treatise treatment.



While the Mixed Range Protobranchia remain relatively stable across the Palaeozoic (Fig.3.20B), in the Treatise dataset clustering intensity drops until the early Carboniferous.

The other two subclades also show differences in clustering intensity between the two treatments. The two curves agree that Heteroconchia reaches its greatest level of clustering by the late Ordovician, suffering a drop across the boundary into the Silurian. This drop is more pronounced in the Treatise treatment (Fig.3.21E). Across the Silurian and Devonian, the intensity of clustering remains stable for the Mixed Range treatment (Fig.3.21B), followed by a drop through the Carboniferous. The Treatise treatment shows a different pattern. Across the Devonian there is a gradual increase in clustering, followed by a gradual decrease up to the End Carboniferous (299 million years ago) (Fig.3.21E). Both acknowledge a sudden increase in clustering across the End Carboniferous boundary where cluster intensity remains for the rest of the Palaeozoic.

The clustering in Pteriomorphia based on the Treatise or the Mixed Range treatments reflects the pattern of Palaeozoic bivalves as a whole (Fig.3.20A&C), although there are some differences between the two treatments (Fig.3.21C&F). The Mixed Range treatment shows a gradual increase in clustering through the Ordovician. The Treatise treatment shows the same temporal pattern of increasing/decreasing cluster intensity as the Mixed Range treatment, but the magnitudes of these changes are different. In the Ordovician both treatments show an increase in clustering intensity. In the Treatise treatment, this increase occurs in the Early Ordovician (Tremadocian-Dapingian) with stability for the remainder of the Ordovician (Fig.3.21F). The Mixed Range treatment shows a more gradual increase across the whole Ordovician (Fig.3.21C).

The fluctuating pattern in pteriomorphian clustering intensity of the Late Ordovician to Late Silurian is shared by both treatments; however, the Devonian drop is more pronounced in the Treatise treatment (Fig.3.21F). As a consequence of this drop, clustering intensity in the Treatise treatment is at its lowest level since the Ordovician origination of the group. There is only a very subtle increase in clustering through the rest of the Palaeozoic. The Mixed Range treatment gives a different story. The drop in clustering intensity in the Devonian is far less pronounced, never reaching the level seen in the Treatise treatment. Therefore the lowest level of clustering in the Mixed Range data set remains at the Ordovician origination of the group. The trend of gradual increase is more strongly seen in the Mixed Range treatment.

### ***3.22 Clustering Intensity Across the Phanerozoic***

#### ***3.23 Principal Points Analysis***

In addition to the nearest neighbour analysis of Foote (1990), principal points analysis was used to describe the clustering of genera in morphospace. The clustering intensity in bivalves based on principal points is illustrated in Figure 3.21. The sample mean standard deviation of the observed taxa is below and outside the range of the

SMSD based on a uniform distribution for PCO axes 4-50. From 50 points onward, the distribution of the data is not distinguishable from random (50-100). Using more than 100 points suggests the data are more spread out than expected by the null uniform distribution. As 4-50 principal points are required to adequately describe the distribution of taxa in the morphospace, this suggests that the total morphospace is highly clustered (Wills *et al.* 2012). This is because the number of points required to describe the space is well below the number of actual points (binomial test  $p$ -value  $< 2.2e-16$ ). This corresponds well with the results of the nearest neighbour analysis and the exploration of subclades through the space. The loss of clustering intensity as principal points are added is expected. As the number of points increases, the accuracy with which the reduced point set can emulate the distribution of the original data increases.

The results of the principal points clustering for the subclades using the morphospace for Bivalvia (Fig.3.24A) using the first 20 PCO axes shows that the morphospace of Bivalvia is highly clustered as only 2 points are required to explain the distribution. The level of clustering in the subclades when considered as part of the Bivalvia morphospace is variable. Protobranchia and Heteroconchia show no clustering (Fig.3.24BC) while Pteriomorphia can be considered clustered using 5-12 principal points. This suggests that the majority of the clustering signal of the Bivalvia is due to clustering in the Pteriomorphia. When the subclades are considered on their own (using either the Mixed Range or Treatise treatments) all are found to be highly clustered themselves.

This shows that when considered within the Bivalvia morphospace, Protobranchia and Heteroconchia are not clustered but there is internal clustering within each subclade. Clustering intensity was plotted through time using the principal points method for stages, epochs and periods but found no evidence for clustering. This result is different to the nearest neighbour analysis that did find some clustering through time. The difference is likely to do with the methods themselves. Principal points seems less sensitive to finding levels of clustering when the number of points in the space is low and may be more appropriate for calculating clustering intensity for the whole space than through time (Fig.3.23).

In summary, the clustering analysis of nearest neighbour and principal point analysis found that the Bivalvia morphospace is highly clustered when considered in its entirety. As time progresses through the Palaeozoic, clustering intensity in Bivalvia increases, but this is primarily the result of clustering of the Pteriomorphia. The Heteroconchia and Protobranchia show less clustering overall than the Pteriomorphia but are also more variable in their levels of clustering through time.

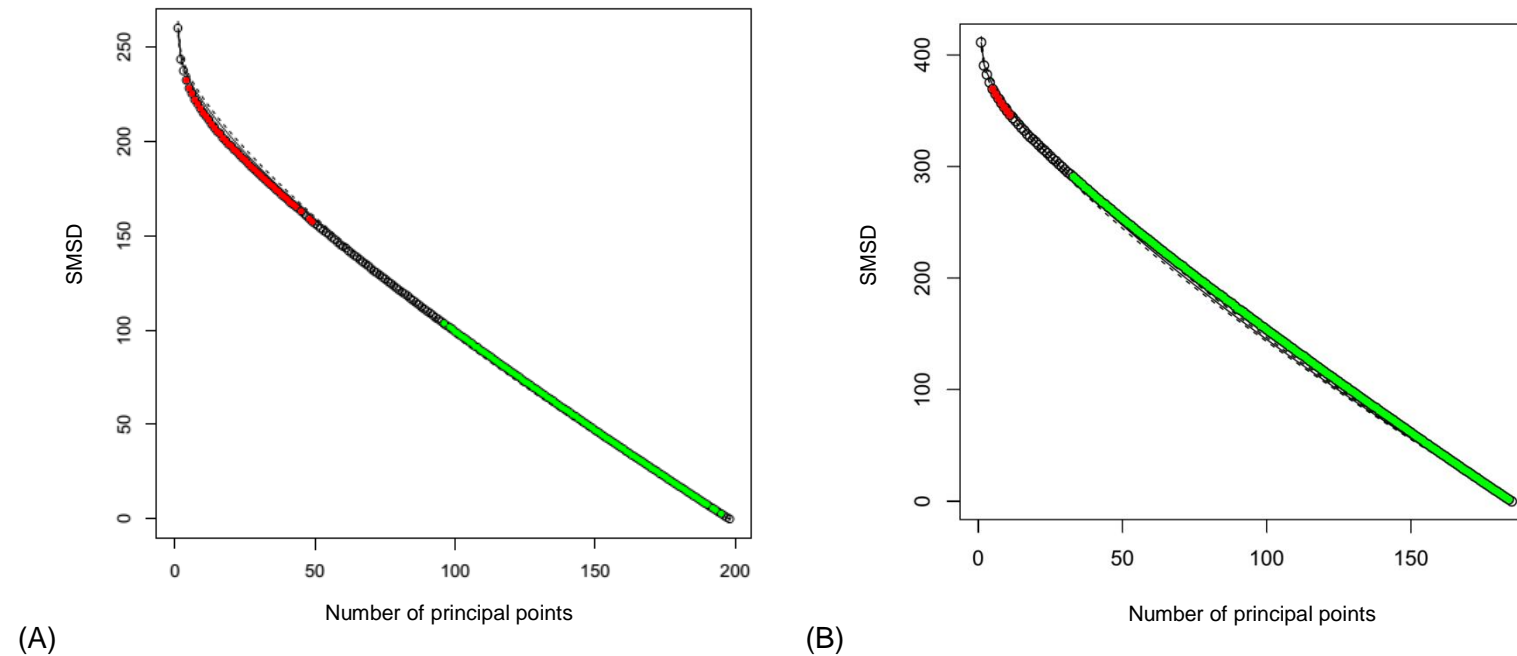


Figure 3.23. Principal points analysis showing the number of principal points needed to fit the distribution of genera within the Palaeozoic bivalve morphospace. The y axis represents the sample mean squared deviation of the measure of performance of k mean clustering for a given set of k principal points. Solid line is the expected SMSD curve for a multivariate homogenous distribution of the same number of points and same spatial bounds as the observed distribution, while the dashed lines are the upper and lower 95% confidence intervals for this null expectation. Dots indicate the empirical SMSD with an increasing number of k principal points. Red are those results which are lower than the lower 95% confidence interval of the null expectation, green are those above the 95% confidence interval. (A) Mixed Range treatment. (B) Treatise treatment.

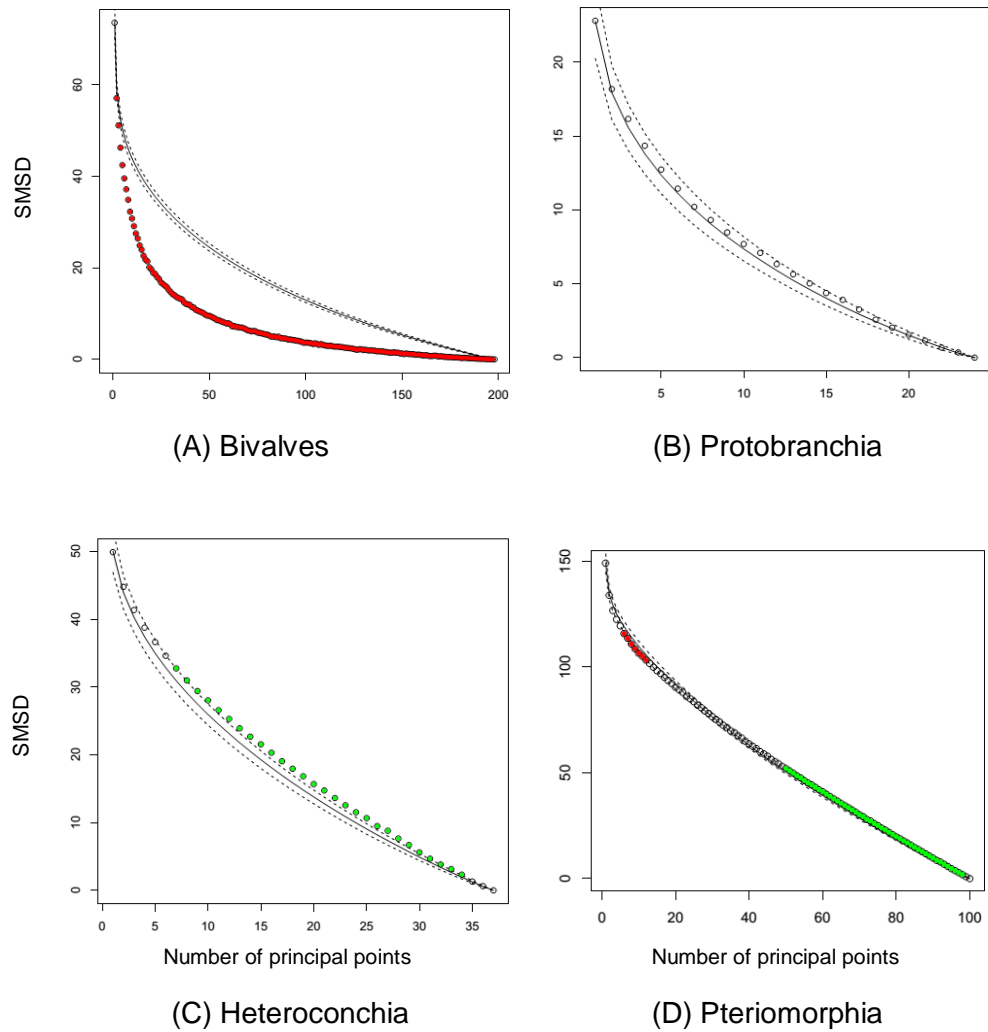


Figure.3.24. Principal points analysis (PPA) of bivalves (A) and the three subclades using the first 20 PCO axes; and (B) Protobranchia; (C) Heteroconchia; (D) and Pteriomorphia, which use all PCO axes. Circles represent the observed sample mean standard deviation (SMSD) of the morphospace for increasing numbers of points. The solid line is the expected SMSD given a multivariate homogenous distribution of the same number of points and special bounds. The dashed lines represent the 95% confidence intervals of the expected distribution. Coloured circles indicate the observed data is found outside the 95% confidence limits of the expected. Red indicates SMSD smaller than expected, green indicates SMSDs larger than expected.

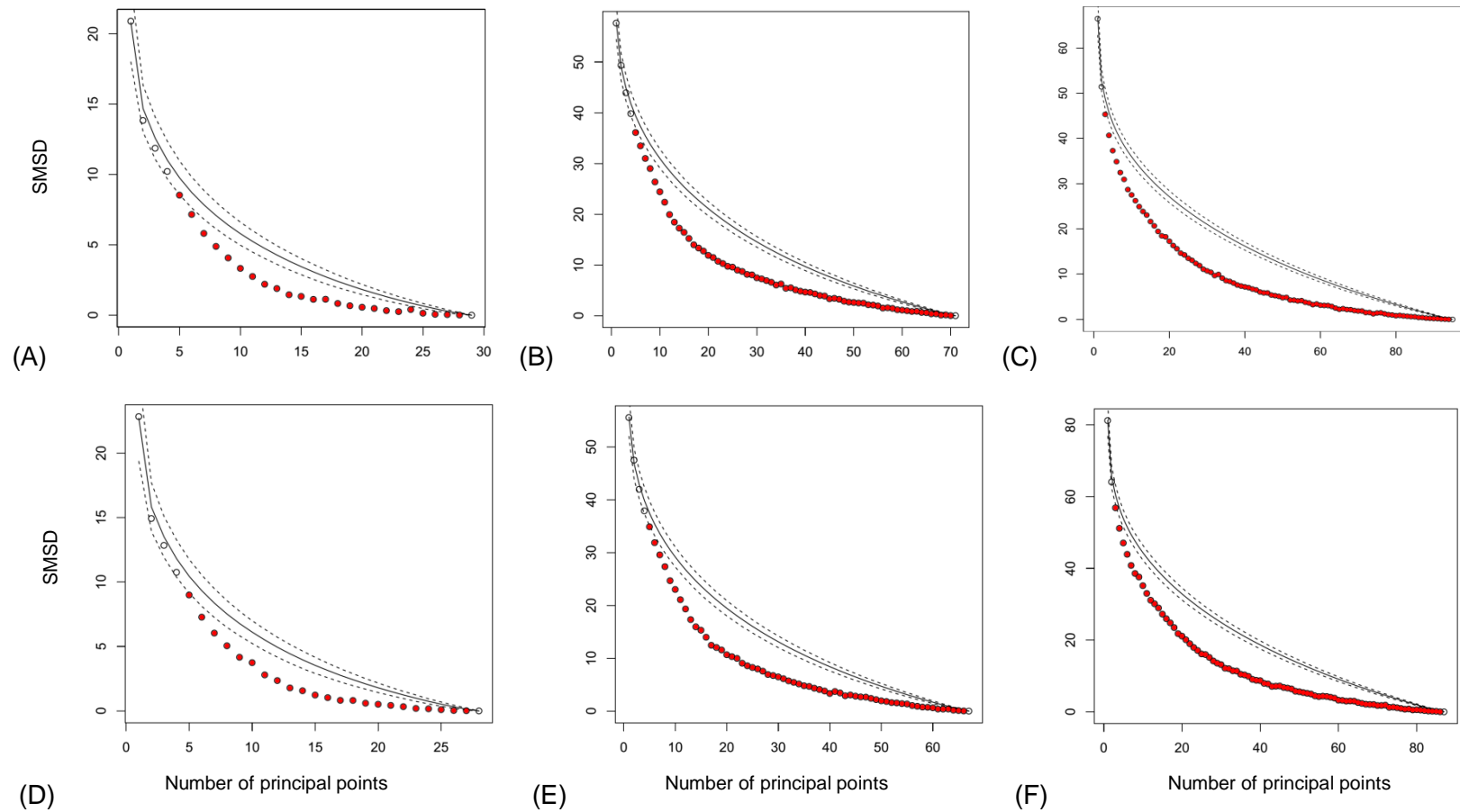


Figure.3.25. PPA of the three subclades using the first 10, 15, and 20 PCO axes respectively. (A) Protobranchia (Mixed Range); (B) Heteroconchia (Mixed Range); (C) Pteriomorpha (Mixed Range); (D) Protobranchia (Treatise); (E) Heteroconchia (Treatise); (F) Pteriomorpha (Treatise). Circles: observed sample mean standard deviation (SMSD). Solid line: expected SMSD. Dashed line: 95% confidence intervals of the expected distribution. Coloured circles indicate the observed data is found outside the 95% confidence limits of the expected. Red indicates SMSD smaller than expected; green indicates SMSDs larger than expected.

### **3.24 Marginal Likelihoods of Disparity Through Time**

The results of the marginal likelihood analysis (Edwards 1992; Friedman 2010) testing for significant jumps in disparity across the Palaeozoic (Mixed Range) are seen in tables.3.3-3.5. The results discussed here are for the Mixed Range treatment only as the Treatise treatment provides the same results. The Treatise treatments data can be found in Appendix.2. Those results that come out as significant are highlighted in bold. The majority of results, regardless of which stratigraphic binning treatment was utilised, did not give likelihood ratio values above the critical value of 8 (Royall 1997). The only exception to this result was across the Cambrian/Ordovician boundary (Table.3.5) which gave a likelihood ratio value of 63465.924. This value is far in excess of the suggested critical values (Royall 1997), and therefore suggest highly significant jumps in disparity across this boundary. The reason for such a large ratio between these two Periods (Cambrian and Ordovician) is due to the large increase in genera with new morphologies across the Cambrian – Ordovician boundary. With the advent of the Ordovician radiation, a whole host of new morphologies evolved, expanding the morphospace in multiple directions (Fig.3.13D). There is a stratigraphic gap between the genera of the Cambrian and Ordovician (Stage 9 and 10 of the Cambrian) where no fossils are found, so there is the chance that some of this jump is due to the paucity of the fossil data in across the interval.

Although there are no significant jumps in disparity across the majority of individual Stages, Epochs and Periods in the Palaeozoic, specific events did have some effect on the disparity of bivalves. Table.3.6 shows the marginal likelihood values for those genera that became extinct vs. those who survived the End Ordovician (Hirnantian – Rhuddanian) and Late Devonian (Frasnian – Famennian) extinction boundaries. The results show ratio values greater than eight for both mass extinction boundaries (i.e., 10.717 and 64.335 for the End Ordovician and Late Devonian respectively). The results suggest that these events could have filtered out those taxa that were less disparate, leading to increasing levels of disparity.

### **3.25 Disparity Through Time**

#### **3.26 Does Disparity Follow a Directional Trend?**

Disparity trajectories (sum of variances) for the three subclades (combined and separate) can be seen in Fig.3.26-3.29. Disparity in the three subclades combined is shown in Figure.3.26. The Mixed Range and Treatise treatments show the same general shape: a gradual, shallow increase in disparity from the Ordovician to the end of the Palaeozoic. This pattern of gradual increase is apparent even when the largest interval binning size (Period) is used. The disparity in the Cambrian was calculated despite the small sample size and taxonomic

Interval of Change		Likelihood ratio			
Interval 1	Interval 2	Bivalvia	Protobranchia	Heteroconchia	Pteriomorphia
-	Tremadocian	-	-	-	-
Tremadocian	Floian	4.085	2.223	0.506	4.649
Floian	Dapingian	1.312	1.174	0.956	1.288
Dapingian	Darriwilian	1.084	0.678	1.021	1.017
Darriwilian	Sandbian	1.054	0.459	1.546	0.799
Sandbian	Katian	1.364	1.997	1.152	1.423
Katian	Hirnantian	0.940	1.366	0.946	0.820
Hirnantian	Rhuddanian	2.070	0.601	0.929	2.372
Rhuddanian	Aeronian	1.031	0.986	1.011	0.988
Aeronian	Telychian	1.008	0.861	1.001	1.025
Telychian	Sheinwoodian	1.011	1.020	1.023	1.138
Sheinwoodian	Homerian	1.075	0.870	0.979	1.078
Homerian	Gorstian	0.916	1.093	1.019	0.903
Gorstian	Ludfordian	1.108	3.266	0.998	0.988
Ludfordian	Pridoli	1.056	1.324	1.002	1.007
Pridoli	Lochkovian	1.569	0.964	1.120	1.172
Lochkovian	Pragian	1.136	1.077	0.854	1.359
Pragian	Emsian	0.963	0.953	0.976	0.994
Emsian	Eifelian	1.085	1.057	1.513	1.004
Eifelian	Givetian	0.960	1.270	0.861	0.888
Givetian	Frasnian	1.830	0.922	1.025	2.127
Frasnian	Famennian	1.463	1.029	0.935	1.120
Famennian	Tournaisian	1.037	1.046	1.061	0.991
Tournaisian	Visean	0.978	1.161	1.007	1.048
Visean	Serpukhovian	1.032	1.205	0.962	1.042
Serpukhovian	Bashkirian	1.025	0.948	1.011	0.980
Bashkirian	Moscovian	1.010	1.017	1.004	1.000
Moscovian	Kasimovian	0.984	0.960	0.991	1.002
Kasimovian	Gzhelian	1.010	1.032	1.016	0.998
Gzhelian	Asselian	1.074	1.014	1.032	0.920
Asselian	Sakmarian	1.018	1.033	0.992	0.996
Sakmarian	Artinskian	0.847	0.949	0.683	0.816
Artinskian	Kungurian	1.125	0.992	0.973	1.159
Kungurian	Roadian	1.021	1.002	1.005	0.991
Roadian	Wordian	0.896	0.918	1.000	0.754
Wordian	Capitanian	1.053	0.967	1.105	0.938
Capitanian	Wuchiapingian	1.219	0.992	1.129	1.060
Wuchiapingian	Changhsingian	1.095	1.006	1.049	1.096

Table.3.3: Marginal likelihood ratios of disparity between successive Stages: Early Ordovician – End Permian (Mixed Range).

Interval of Change		Likelihood ratio			
Interval 1	Interval 2	Bivalvia	Protobranchia	Heteroconchia	Pteriomorphia
-	L.Ordovician	-	-	-	-
L.Ordovician	M.Ordovician	1.556	0.803	0.956	1.387
M.Ordovician	U.Ordovician	1.504	0.751	2.148	1.091
U.Ordovician	Llandovery	2.042	0.626	0.904	1.859
Llandovery	Wenlock	1.066	0.926	0.995	1.259
Wenlock	Ludlow	1.037	3.622	1.033	0.905
Ludlow	Pridoli	1.036	1.398	0.990	0.987
Pridoli	L.Devonian	1.710	1.002	1.133	1.535
L.Devonian	M.Devonian	1.165	0.957	1.190	1.101
M.Devonian	U.Devonian	2.849	1.158	1.031	2.552
U.Devonian	Mississippian	1.172	1.239	0.866	1.091
Mississippian	Pennsylvanian	0.967	1.179	1.009	0.889
Pennsylvanian	Cisuralian	1.049	1.006	0.704	0.847
Cisuralian	Guadalupian	0.965	0.977	1.053	0.676
Guadalupian	Lopingian	1.337	0.891	1.163	1.264

Table.3.4: Marginal likelihood ratios of disparity between successive Epochs: Early Ordovician – End Permian (Mixed Range).

Interval of Change		Likelihood ratio			
Interval 1	Interval 2	Bivalvia	Protobranchia	Heteroconchia	Pteriomorphia
-	Cambrian	-	-	-	-
Cambrian	Ordovician	<b>63465.924</b>	-	-	-
Ordovician	Silurian	4.107	0.951	1.001	2.307
Silurian	Devonian	5.771	1.617	2.022	4.591
Devonian	Carboniferous	1.510	1.420	0.764	1.170
Carboniferous	Permian	1.015	1.150	0.698	0.559

Table.3.5: Marginal likelihood ratios of disparity between successive Periods: Early Ordovician – End Permian (Mixed Range). Significant values highlighted in bold.

uncertainty of genera residing within the period. For all statistics relating to the disparity in the Palaeozoic, the Cambrian was excluded because the paucity of data. Unless specifically stated all inferences in the Palaeozoic refers to the Ordovician to Permian periods. The level of disparity in the Cambrian is highly variable across the two treatments and the three interval binning sizes (Fig.3.26). This is the result of the paucity of taxa in individual bins; reaching a maximum value of three genera in Cambrian Stage 4 (Mixed Range treatment).



	Ordovician Extinction		Devonian Extinction	
	Extinct	Survive	Extinct	Survive
Disparity	215.750	242.765	225.481	261.007
Standard error	11.570	3.667	15.167	3.034
Likelihood ratio	<b>10.717</b>		<b>64.335</b>	

Table.3.6: Marginal likelihood ratio for the change in disparity across the End Ordovician (Hirnantian - Rhuddanian) and Late Devonian (Frasnian - Famennian) mass extinction boundaries. Disparity: sum of variances of genera that go extinct at the boundary or survive it. Standard error from 1000 bootstrap replicates. Likelihood ratio: marginal likelihood ratio across the boundary. Significant values highlighted in bold.

The disparity trajectory of Protobranchia is found in Fig.3.27. The Mixed Range and Treatise treatments show a relatively similar trend in disparity up to the end of the Devonian, but there are some noticeable differences. In the Treatise treatment, the seemingly rapid increase in disparity in the Late Silurian is not present; instead this rise is more gradual. There is no evidence; however, that the jump is significant due to the low marginal likelihood ratios over these intervals (Tables 3.3-3.5) and overlapping standard error bars. Both the Mixed Range and Treatise treatments pick up a rise in disparity through the Ordovician (Stage and Epoch). The two treatments disagree regarding the direction of disparity change across the Silurian (Epoch); with the Mixed Range data showing an increase and the Treatise data showing a decrease (Fig.3.27AB, Period). From this point on, the two patterns concur on an increase in disparity across the Devonian. After this stage, the Mixed Range data shows a gradually increasing trend in disparity, whereas disparity in the Treatise treatment remains stationary. This difference is due to the lack of taxa present in these intervals according to the Treatise treatment compared to the Mixed Range treatment.

The disparity trajectory of Heteroconchia is shown in Figure.3.27. Both the Mixed Range and Treatise treatments agree in general shape across Stage, Epoch and Period interval binning. The curve in all treatments of the Heteroconchia remains relatively static across the whole Palaeozoic, with only minor (and non-significant) fluctuations. The biggest changes are ones of increase in the Ordovician; however, this jump is not significant in likelihood ratio tests (Tables 3.3-3.5). This relatively flat pattern of disparity is backed up by the stable clustering through time for Heteroconchia (Fig.22B, 23B and 24B), and lack of travel of the clade through the morphospace (Fig.24-26). After an initial exploration of morphospace in the Ordovician, the majority of new taxa occupy areas within the limits set out by preceding taxa. The pteriomorphs are the clade with the disparity pattern most similar to that found in bivalves as a whole (Fig.3.29). Some subtle differences can be seen between the disparity patterns of the Mixed Range and Treatise treatments. In the Mixed Range treatment, a gradual increase is observed (Fig.3.26),

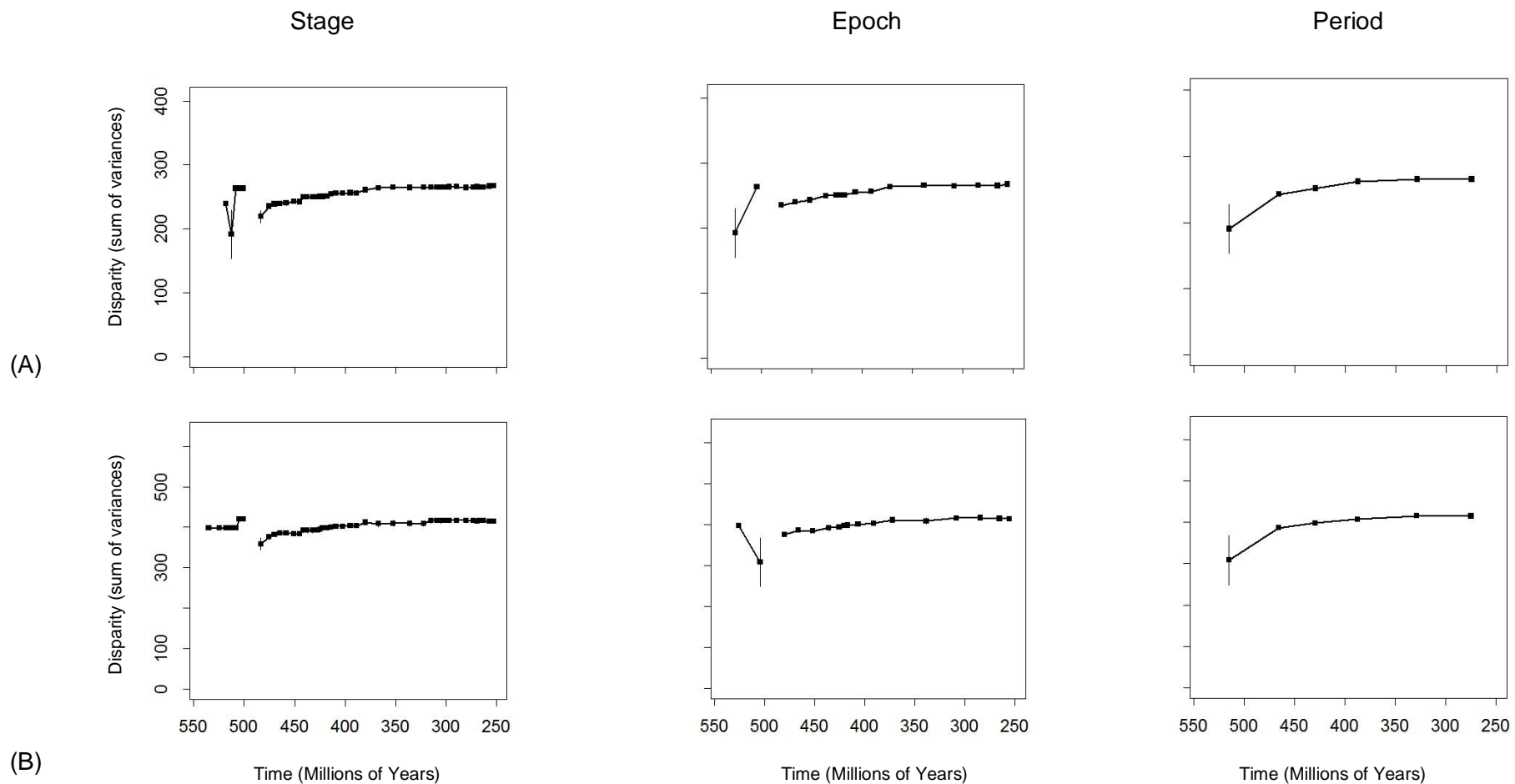


Figure.3.26: Disparity (sum of variances) trajectories for Palaeozoic bivalves using different stratigraphic binning regimes. The three regimes are Stage, Epoch and Period. Disparity is plotted at the midpoint of each interval. (A) Mixed Range treatment. (B) Treatise treatment. Black dots: observed mean disparity based on 1000 bootstrap replicates. Standard error bars are based on 1000 bootstrap replicates.

whilst in the Treatise treatment the pattern at all interval binning sizes is relatively flat across the Palaeozoic (Fig.3.29).

Statistical evidence for the gradual increase in disparity across the Palaeozoic (Fig.3.26) can be found in Tables 3.4-3.7. Spearman Rank correlation coefficients (Table 3.4) between disparity (sum of variances) and time are reported as highly significant (Spearman's  $\rho$ :  $>0.8$ ,  $P$  value  $< 0.01$ ) positive correlation in the majority of Stage and Epoch interval treatments for both the Mixed Range and Treatise treatments. Period level binning for bivalves and Protobranchia are also highly significant and positively correlated, but only in the Mixed Range treatment. The period level interval binning with the Treatise treatments of all bivalves and each subclade – in addition to the Mixed Range treatments in Heteroconchia and Pteriomorphia - show positive but non-significant correlations between disparity and time. This is not entirely unexpected. As time interval binning increases, the number of genera included in each bin increases. Inevitably this increases the probability that the taxa sampled will have explored more extremes of the morphospace thereby approaching the extent of the space defined by all taxa. This means that only radical changes in morphospace occupation between successive time bins are likely to be detectable.

The analysis of the Chapter 1 (Hughes *et al.* 2013) found that the majority of extinction ended clades, and many extant clades had top-heavy disparity trajectories. Isolating the Palaeozoic effectively creates an artificial extinction event at the End Permian boundary which could favour finding a top-heavy (increasing disparity) profile; however, as the results of Chapter 1 show, bivalves show a top-heavy pattern even when the whole Phanerozoic is taken into account. To obtain these top-heavy trajectories., Foote (1993b) and Wesley-Hunt (Wesley-Hunt 2005) suggested two approaches that could result in a top-heavy trajectory of disparity: (1) constrained morphology continuing to rise as diversity increases, or (2) concordant growth in diversity and disparity with no limit on morphological evolution. Either process would result in a linear increase of disparity with time (and diversity). I therefore tested whether the increasing disparity of bivalves also followed a linear trend.

Linear regressions of each treatment (Table 3.7) suggest a significant correlation between time and disparity ( $R^2 > 0.8$ ,  $P$  values'  $< 0.03$ ). The change in disparity with each successive interval based on marginal likelihood ratios (Tables 3.3-3.5) found no evidence for punctuated changes in disparity over the Palaeozoic (Ordovician-Permian). Looking at the first differences in disparity between successive intervals of time provides evidence of a decreasing net increase in disparity with each interval. Most treatments find this negative relationship as non-significant (Table 3.7: Spearman's  $\rho < 0$ ,  $P$  values  $> 0.05$ ). At the smallest interval bin size (Stage) both the Mixed Range and Treatise treatments find a significant relationship between the decreasing net increase in disparity and time (for

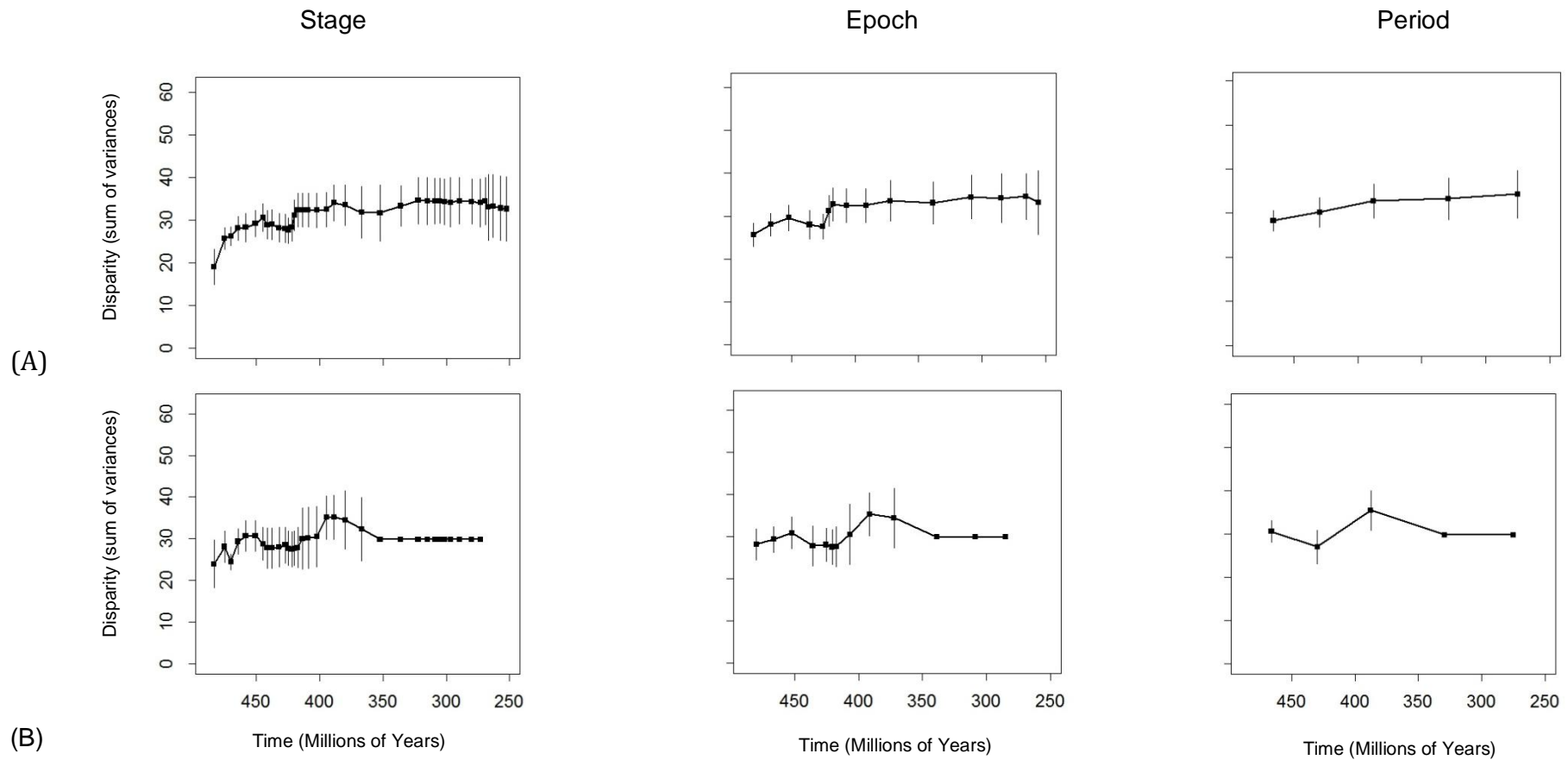


Figure.3.27: Disparity (sum of variances) trajectories for Palaeozoic Protobranchia using different stratigraphic binning regimes. The three regimes are Stage, Epoch and Period. Disparity is plotted at the midpoint of each interval. (A) Mixed Range treatment. (B) Treatise treatment. Black dots: observed mean disparity based on 1000 bootstrap replicates. Standard error bars are based on 1000 bootstrap replicates.

Bivalvia), but time itself does not explain much of this drop (Mixed Range:  $R^2 = 0.1401$ , P value = 0.02246; Treatise:  $R^2 = 0.1377$ , P value = 0.02373).

So far, the evidence points to an increase in bivalve disparity across the Palaeozoic (Ordovician – Permian) showing no strong evidence for punctuated change and even some preliminary evidence for a drop in the rate at which disparity increases with each interval. The pattern seems most similar to an evolutionary model describing an Ornstein-Uhlenbeck process (OU), i.e., initially rapid increase of a trait or parameter (disparity in this case) followed by a reduction in rate as some local optimum (or constraint) is reached. The original use of the OU model, like the other models (GRW, URW and STA) was to simulate the evolutionary trajectory of individual traits. The trajectory of disparity is a property of the clade of interest, and will likely relate to changes that effect the clade on a broad scale for example changes in common environments, predators, food type or habitat use or more likely a complex combination of these.

Table 3.8 shows the results of fitting several evolutionary models on disparity across the Palaeozoic (Hunt 2006; Hunt *et al.* 2008; Hunt 2008). Of the four evolutionary models: GRW, URW, STA and OU, the last model fits best in the majority of treatments. There is a difference between the model chosen for the Mixed Range and Treatise treatments the Protobranchia; GRW (directional) was most likely (43%) followed by URW (random exploration through the space) (35%). In the period interval binning data, the model that best fits the bivalves and each subclade is the random walk model.

The Ornstein-Uhlenbeck algorithm currently does not implement the ancestor-descendant (AD) parameterisation, which considers the morphological differences of ancestor-descendent pair of populations (disparity in each time interval) as separate observations. Each observation is expected to be normally distributed with means and variances that are functions of model parameters and elapsed time (Hunt 2006; Hunt *et al.* 2008; Hunt 2008). All models implement the joint parameterisation which consider all disparity values jointly, and to have come from a single draw of a multivariate normal distribution. The vector of means and covariance matrix are functions of the model parameters and age model (Hunt 2006; Hunt *et al.* 2008; Hunt 2008).

By removing the OU model from the analysis, the models were fitted to the data again to see whether hanging the parameterisation would affect the results. Table 3.9 and 3.10 show the models that best fit the data once the OU model was removed using both the AD and joint parameterisations. There is far less agreement which model fits the data when the Ornstein-Uhlenbeck model is removed. In bivalves, the directional model of evolution comes out on top (Akaike weights of >0.8) at the level of stage, but as time intervals increase, the URW model becomes the selected model. For the subclades, the URW is the most chosen model, but the weights between the directional and random walk (GRW and URW) are closer together (e.g. 0.491 and 0.509) suggesting that the algorithms are finding it rather difficult to pick one model over the other (Table 3.9-10).

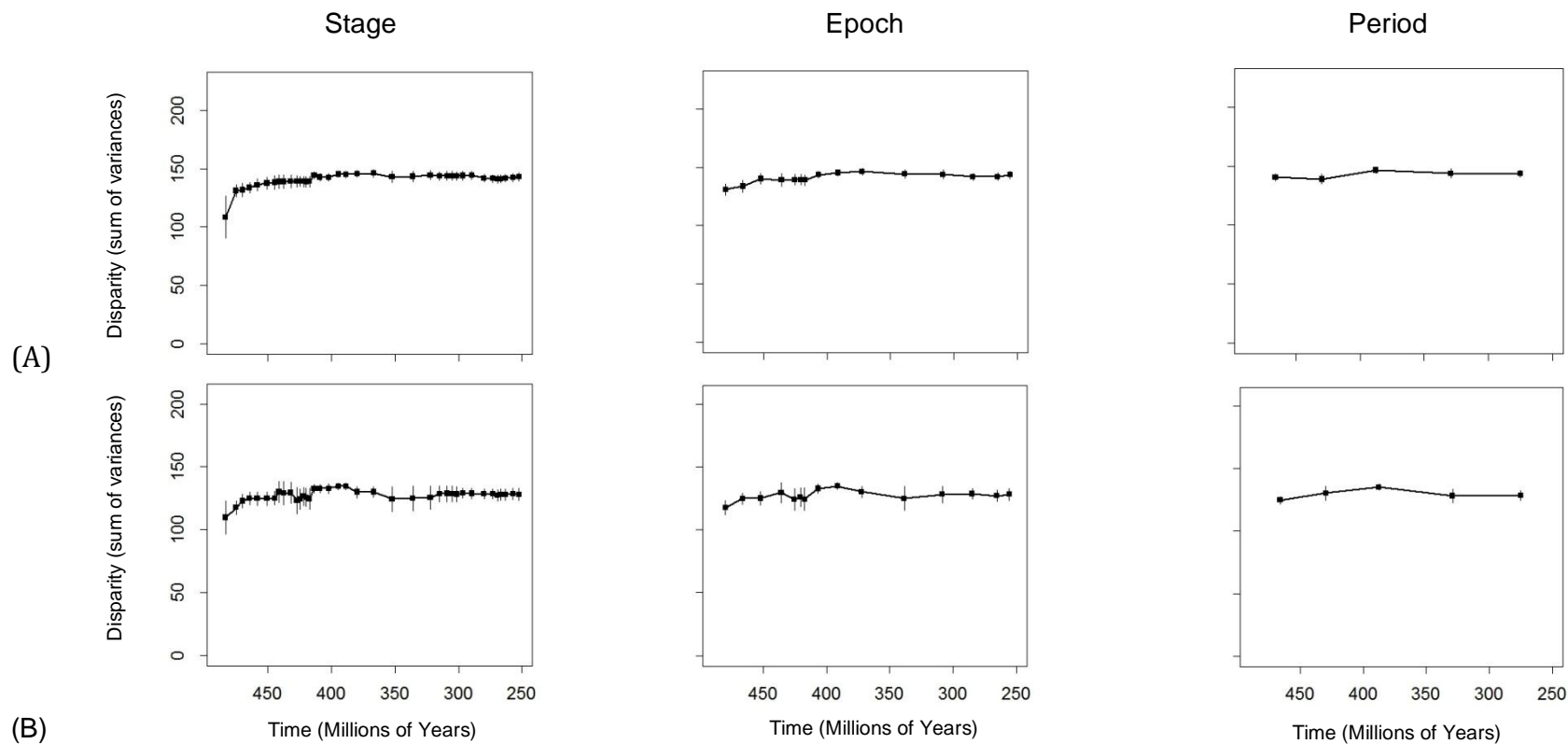


Figure.3.28: Disparity (sum of variances) trajectories for Palaeozoic Heteroconchia using different stratigraphic binning regimes. The three regimes are Stage, Epoch and Period. Disparity is plotted at the midpoint of each interval. (A) Mixed Range treatment. (B) Treatise treatment. Black dots: observed mean disparity based on 1000 bootstrap replicates. Standard error bars are based on 1000 bootstrap replicates.

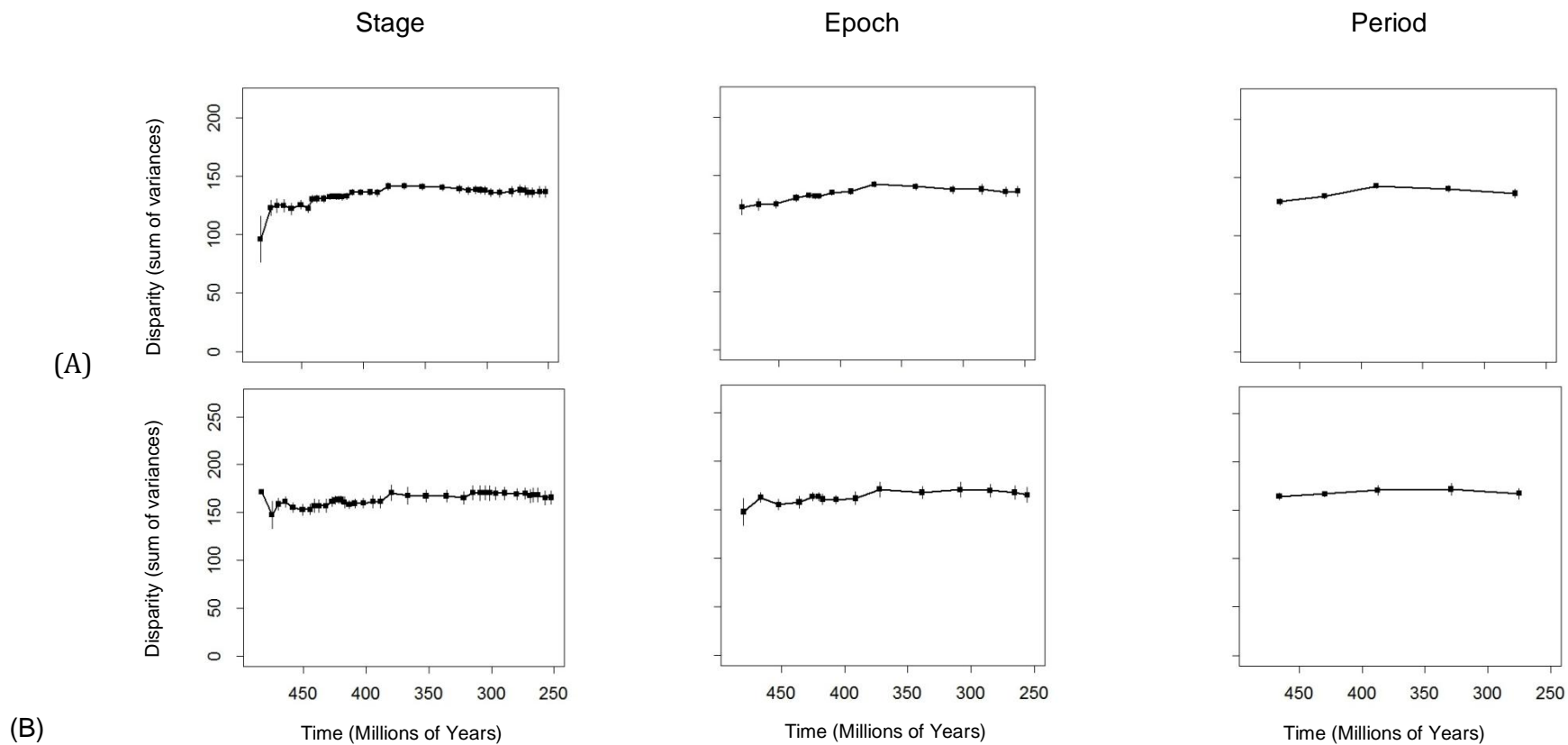


Figure.3.29: Disparity (sum of variances) trajectories for Palaeozoic Pteriomorphia using different stratigraphic binning regimes. The three regimes are Stage, Epoch and Period. Disparity is plotted at the midpoint of each interval. (A) Mixed Range treatment. (B) Treatise treatment. Black dots: observed mean disparity based on 1000 bootstrap replicates. Standard error bars are based on 1000 bootstrap replicates.

"Raw" disparity		STAGE		Epoch		PERIOD	
Clade	Treatment	$\rho$	P-value	$\rho$	P-value	$P$	P-value
Bivalves	Mixed Range	0.972	<b>2.20E-16</b>	0.982	<b>2.20E-16</b>	1	<b>1.67E-02</b>
	Treatise	0.917	<b>2.20E-16</b>	0.961	<b>2.20E-16</b>	1	0.083
Protobranchia	Mixed Range	0.825	<b>3.23E-08</b>	0.911	<b>2.20E-16</b>	1	<b>1.67E-02</b>
	Treatise	0.442	<b>0.01007</b>	0.387	0.0787	-0.154	0.805
Heteroconchia	Mixed Range	0.64	<b>2.33E-05</b>	0.696	<b>5.18E-03</b>	0.5	0.45
	Treatise	0.336	<b>0.03971</b>	0.432	0.109	0.2	0.783
Pteriormorphia	Mixed Range	0.765	<b>2.78E-07</b>	0.85	<b>4.10E-05</b>	0.6	3.50E-01
	Treatise	0.661	<b>1.06E-05</b>	0.764	<b>1.39E-03</b>	0.7	2.33E-01

First differences		STAGE		Epoch		PERIOD	
Clade	Treatment	$\rho$	P-value	$P$	P-value	$P$	P-value
Bivalves	Mixed Range	-0.271	0.105	-0.402	0.155	-0.8	0.333
	Treatise	-0.315	0.058	-0.451	0.108	-1	0.333
Protobranchia	Mixed Range	-0.373	<b>0.023</b>	-0.323	0.26	-0.6	0.417
	Treatise	-0.195	0.285	-0.291	0.359	0	1
Heteroconchia	Mixed Range	-0.331	<b>0.046</b>	-0.38	0.181	0	1
	Treatise	-0.222	0.187	-0.235	0.418	-0.8	0.333
Pteriormorphia	Mixed Range	-0.244	0.145	-0.442	0.116	-0.8	0.333
	Treatise	-0.116	0.494	-0.279	0.333	-0.8	0.333

Table 3.7: Results of Spearman Rank correlation coefficient and (sum of variances) vs. time. The “raw” disparity values refer to the mean disparity values per interval outputted from the disparity analysis. Significant values are highlighted in bold.



Interval bin	Treatment	Model (Joint)	Bivalves			Protobranchia			Heteroconchia			Pteriomorphia		
			log(L)	AICc	AW	log(L)	AICc	AW	log(L)	AICc	AW	log(L)	AICc	AW
Stage	Mixed Range	Directional	-89.965	186.636	0	-75.942	158.589	0.005	-97.068	200.842	0	-106.627	219.96	0
		Random walk	-93.453	191.248	0	-77.253	158.85	0.004	-98.496	201.335	0	-107.81	219.962	0
		Stasis	-145.348	295.039	0	-99.92	204.183	0	-124.25	252.843	0	-133.609	271.561	0
		Ornstein-Uhlenbeck	-79.662	168.535	<b>1</b>	-69.318	147.845	<b>0.991</b>	-71.97	153.152	<b>1</b>	-90.618	190.449	<b>1</b>
	Treatise	Directional	-98.865	204.436	0	-103.692	214.087	<b>0.431</b>	-91.797	190.3	0.001	-132.521	271.749	0
		Random walk	-101.949	208.241	0	-104.82	213.983	<b>0.454</b>	-92.408	189.158	0.002	-133.536	271.414	0
		Stasis	-154.484	313.31	0	-141.981	288.306	0	-109.371	223.085	0	-152.085	308.513	0
		Ornstein-Uhlenbeck	-89.91	189.033	<b>0.999</b>	-103.753	216.719	<b>0.116</b>	-83.696	176.603	<b>0.997</b>	-102.87	214.952	<b>1</b>
Epoch	Mixed Range	Directional	-32.581	73.343	0.073	-29.275	66.733	0.1	-32.962	74.105	0.044	-33.091	74.364	<b>0.171</b>
		Random walk	-36.511	78.022	0.007	-30.712	66.423	<b>0.117</b>	-33.826	72.652	0.09	-34.109	73.218	<b>0.304</b>
		Stasis	-55.83	116.658	0	-36.148	77.296	0.001	-42.458	89.917	0	-47.156	99.312	0
		Ornstein-Uhlenbeck	-28.14	68.279	<b>0.92</b>	-25.314	62.629	<b>0.782</b>	-28.068	68.135	<b>0.866</b>	-30.061	72.123	<b>0.525</b>
	Treatise	Directional	-84.663	177.508	<b>0.104</b>	-43.601	95.385	<b>0.386</b>	-40.366	88.913	0.043	-44.351	96.884	0.085
		Random walk	-85.203	175.406	0	-44.793	94.587	<b>0.575</b>	-40.616	86.232	<b>0.163</b>	-44.743	94.486	<b>0.283</b>
		Stasis	-90.44	185.881	0	-56.561	118.123	0	-41.96	88.92	0.042	-47.984	100.968	0.011
		Ornstein-Uhlenbeck	-35.049	82.098	<b>0.864</b>	-43.996	99.993	0.039	-35.583	83.167	<b>0.753</b>	-40.459	92.917	<b>0.62</b>
Period	Mixed Range	Directional	-11.748	53.497	0	-5.832	41.665	0.001	-11.795	53.589	0	-12.807	55.614	0
		Random walk	-13.375	36.749	<b>0.991</b>	-9.126	28.252	<b>0.868</b>	-11.857	33.714	<b>0.543</b>	-13.009	36.017	<b>0.857</b>
		Stasis	-18.093	46.186	0.009	-11.015	32.03	<b>0.131</b>	-12.027	34.055	<b>0.457</b>	-14.801	39.603	<b>0.143</b>
		Ornstein-Uhlenbeck	-7.314	NA	0	-5.053	NA	0	-10.09	NA	0	-10.507	NA	0
	Treatise	Directional	-12.85	55.7	0	-15.715	61.43	0	-12.839	55.677	0	-11.128	52.256	0
		Random walk	-14.701	39.403	<b>0.988</b>	-16.048	42.096	<b>0.833</b>	-12.91	35.821	<b>0.59</b>	-11.239	32.477	<b>0.656</b>
		Stasis	-19.084	48.168	0.012	-17.653	45.306	<b>0.167</b>	-13.272	36.545	<b>0.41</b>	-11.813	33.625	<b>0.344</b>
		Ornstein-Uhlenbeck	-7.187	NA	0	-15.11	NA	0	-10.354	NA	0	-9.093	NA	0

Table.3.8: The fit of each evolutionary model to the data (bivalves and each subclade) based on likelihood analysis. For each type of interval binning under the two treatments, one of four evolutionary models (directional, random stasis, and Ornstein-Uhlenbeck) was fit using maximum likelihood. Log likelihood: log(L). Akaike information criterion corrected for sample size: AICc. Akaike weight: AW.. All results refer to joint parameterisation only. Strongly supported results are highlighted in bold.

Interval bin	Treatment	Model	Bivalves						Protobranchia					
			AD			Joint			AD			Joint		
			log(L)	AICc	AW	log(L)	AICc	AW	log(L)	AICc	AW	log(L)	AICc	AW
Stage	Mixed Range	Directional	-89.243	182.838	<b>0.906</b>	-89.965	186.636	<b>0.909</b>	-72.015	148.383	<b>0.408</b>	-75.942	158.589	<b>0.533</b>
		Random walk	-92.632	187.378	0.094	-93.453	191.248	0.091	-72.762	147.638	<b>0.592</b>	-77.253	158.85	<b>0.467</b>
		Stasis	-135.872	276.098	0	-145.348	295.039	0	-88.988	182.328	0	-99.92	204.183	0
	Treatise	Directional	-97.301	198.954	<b>0.844</b>	-98.865	204.436	<b>0.87</b>	-100.17	204.693	<b>0.262</b>	-103.691	214.087	<b>0.487</b>
		Random walk	-100.104	202.323	<b>0.156</b>	-101.949	208.241	<b>0.13</b>	-100.255	202.624	<b>0.738</b>	-104.82	213.983	<b>0.513</b>
		Stasis	-145.588	295.528	0	-154.484	313.31	0	-138.792	281.936	0	-141.981	288.306	0
Epoch	Mixed Range	Directional	-32.522	70.136	<b>0.923</b>	-32.581	73.343	<b>0.912</b>	-26.975	59.04	<b>0.251</b>	-29.275	66.733	<b>0.46</b>
		Random walk	-36.387	75.107	0.077	-36.511	78.022	0.088	-27.262	56.858	<b>0.747</b>	-30.712	66.423	<b>0.537</b>
		Stasis	-50.313	105.716	0	-55.829	116.658	0	-31.927	68.944	0.002	-36.148	77.296	0.002
	Treatise	Directional	-83.846	172.783	0	-84.663	177.508	<b>0.258</b>	-42.093	89.277	<b>0.348</b>	-43.601	95.385	<b>0.402</b>
		Random walk	-84.386	171.105	0	-85.203	175.406	<b>0.738</b>	-42.845	88.023	<b>0.652</b>	-44.793	94.587	<b>0.598</b>
		Stasis	-54.972	115.035	<b>1</b>	-90.44	185.881	0.004	-53.24	111.57	0	-56.561	118.123	0
Period	Mixed Range	Directional	-12.191	40.383	0.012	-11.748	53.497	0	-6.1082	28.215	0.017	-5.832	41.665	0.001
		Random walk	-13.817	31.635	<b>0.98</b>	-13.376	36.749	<b>0.991</b>	-8.062	20.123	<b>0.978</b>	-9.126	28.252	<b>0.868</b>
		Stasis	-12.6419	41.284	0.008	-18.093	46.186	0.009	-7.395	30.79	0.005	-11.015	32.03	<b>0.131</b>
	Treatise	Directional	-12.649	41.299	0.015	-12.85	55.7	0	-14.871	45.742	0.003	-15.715	61.43	0
		Random walk	-14.496	32.992	<b>0.978</b>	-14.701	39.403	<b>0.987</b>	-15.132	34.263	<b>0.991</b>	-16.048	42.096	<b>0.833</b>
		Stasis	-13.466	42.932	0.007	-19.084	48.168	0.012	-14.348	44.697	0.005	-17.653	45.306	<b>0.167</b>

Table.3.9: The fit of each evolutionary model to the data (bivalves and Protobranchia) based on likelihood analysis. For each type of interval binning for the two treatments, one of three evolutionary models (directional, random stasis) was fit using maximum likelihood. Log likelihood: log(L). Akaike information criterion corrected for sample size: AICc. Akaike weight: AW. Two parameterisations were possible: ancestor-descendant (AD), and joint. Both are reported. Strongly supported results are highlighted in bold.

Interval bin	Treatment	Model	Heteroconchia						Pteriomorphia					
			log(L)	AICc	AW	log(L)	AICc	AW	log(L)	AICc	AW	log(L)	AICc	AW
Stage	Mixed Range	Directional	-95.549	195.451	<b>0.536</b>	-97.068	200.842	<b>0.561</b>	-105.152	214.657	<b>0.491</b>	-106.627	219.96	<b>0.5</b>
		Random walk	-96.826	195.766	<b>0.457</b>	-98.496	201.335	<b>0.439</b>	-106.234	214.582	<b>0.509</b>	-107.81	219.962	<b>0.5</b>
		Stasis	-99.879	204.11	0.007	-124.25	252.843	0	-114.353	233.058	0	-133.609	271.561	0
	Treatise	Directional	-88.192	180.736	<b>0.265</b>	-91.797	190.3	<b>0.361</b>	-129.119	262.59	0	-132.521	271.749	<b>0.458</b>
		Random walk	-88.291	178.697	<b>0.735</b>	-92.408	189.158	<b>0.639</b>	-130.085	262.285	0	-133.536	271.414	<b>0.542</b>
		Stasis	-98.367	201.088	0	-109.371	223.085	0	-117.879	240.112	1	-152.085	308.513	0
Epoch	Mixed Range	Directional	-31.777	68.643	<b>0.332</b>	-32.962	74.105	<b>0.326</b>	-32.145	69.38	<b>0.367</b>	-33.091	74.364	<b>0.36</b>
		Random walk	-32.462	67.257	<b>0.664</b>	-33.826	72.652	<b>0.674</b>	-32.977	68.288	<b>0.633</b>	-34.109	73.218	<b>0.64</b>
		Stasis	-36.22	77.532	0.004	-42.458	89.917	0	-42.253	89.597	0	-47.156	99.312	0
	Treatise	Directional	-39.742	84.575	0.018	-40.367	88.913	<b>0.172</b>	-43.095	91.28	0.095	-44.351	96.884	<b>0.225</b>
		Random walk	-39.917	82.166	0.061	-40.616	86.232	<b>0.657</b>	-43.405	89.144	<b>0.277</b>	-44.743	94.487	<b>0.746</b>
		Stasis	-35.8161	76.723	<b>0.921</b>	-41.96	88.92	<b>0.171</b>	-41.21	87.51	<b>0.627</b>	-47.984	100.968	0.029
Period	Mixed Range	Directional	-11.484	38.968	0.003	-11.795	53.589	0	-12.464	40.928	0.003	-12.807	55.614	0
		Random walk	-11.538	27.075	<b>0.983</b>	-11.857	33.714	<b>0.543</b>	-12.662	29.324	<b>0.977</b>	-13.009	36.017	<b>0.857</b>
		Stasis	-9.743	35.487	0.015	-12.027	34.055	<b>0.457</b>	-10.545	37.091	0.02	-14.801	39.603	<b>0.143</b>
	Treatise	Directional	-12.107	40.213	0.003	-12.839	55.677	0	-10.188	36.376	0.003	-11.128	52.256	0
		Random walk	-12.173	28.347	<b>0.967</b>	-12.911	35.821	<b>0.59</b>	-10.263	24.525	<b>0.982</b>	-11.239	32.477	<b>0.64</b>
		Stasis	-9.644	35.289	0.03	-13.272	36.545	<b>0.41</b>	-8.451	32.903	0.015	-11.813	33.625	<b>0.36</b>

Table.3.10: The fit of each evolutionary model to the data (Heteroconchia and Pteriomorphia) based on likelihood analysis. For each type of interval binning for the two treatments, one of three evolutionary models (directional, random, and stasis) was fit using maximum likelihood. Log likelihood: log(L). Akaike information criterion corrected for sample size: AICc. Akaike weight: AW. Two parameterisations were possible: ancestor-descendant (AD), and joint. Both are reported. Strongly supported results are highlighted in bold.

The model with the lowest level of support is the stasis model. As the interval binning increases from Stage to Period, the random walk model tends to be picked over the others models (whether the OU is included or not). Overall, the results of the model fitting suggest that the disparity trajectories of the Bivalvia and the subclades Protobranchia, Heteroconchia and Pteriomorphia can best be described by the Ornstein-Uhlenbeck process of evolution when it is included, and directional or random when it is not.

### **3.27 Partial Disparity Across the Palaeozoic**

The contribution of each subclade to the disparity of bivalves across the Palaeozoic is shown in Figures 3.30 and 3.31. The major trend in each partial disparity plot is one of increase across the Palaeozoic, as seen in Fig.3.26. The general trends of each subclade are effectively the same across each treatment; however, the changes in the Treatise treatment are more pronounced. Across the Palaeozoic, Protobranchia contribute the least to total bivalve disparity, followed by the Heteroconchia, and finally the Pteriomorphia which contributed the most. The contribution of each subclade is comparatively dynamic across time, with the increase in one resulting in the relative drop by another. The main difference between the Treatise and Mixed Range treatments is the change in contribution between the three subclades. The Mixed Range treatment (Fig.3.30) shows a more gradual change in the contributions made by each subclade while the Treatise treatment the changes are more rapid and tend to be larger (Fig.3.31).

Protobranchia contribute the most to the disparity curve in the Early-Middle Ordovician (488.3-460.9 million years ago) (Fig.3.30 and 3.31). Their contribution steadily declines from this point onward until the end of the Palaeozoic. The contribution by the Heteroconchia is slightly different between the Mixed Range and Treatise treatments. In the Mixed Range treatment the contribution by Heteroconchia to disparity is stable through the Ordovician and Silurian (Fig.3.30), while in the Treatise treatment the contribution of Heteroconchia drops in the Middle Ordovician (Fig.3.31). Through the Devonian both treatments show an increase in the contribution made by Heteroconchia, but the relative magnitude of increase is greater in the Treatise treatment (Fig.3.31). Through the Carboniferous the Mixed Range treatment indicates that the Heteroconchia stabilise in their contribution, no longer fluctuating up to the End Permian. (Fig.3.30), The Treatise treatment shows an additional increase in the contribution by the Heteroconchia at the beginning of the Permian (Asselian: 299 million years ago), accompanied by a drop in the Pteriomorphia and Protobranchia.

The Mixed Range treatment shows the Pteriomorphia go through a minor decline in the Early Ordovician as Protobranchia are increasing in disparity. For the result of the Palaeozoic Pteriomorphia increase gradually up to the End Permian boundary, becoming the largest contributor to bivalve disparity (Fig.3.30). The Treatise treatment gives an account that mirrors the Mixed Range except for the comparatively larger decline in their

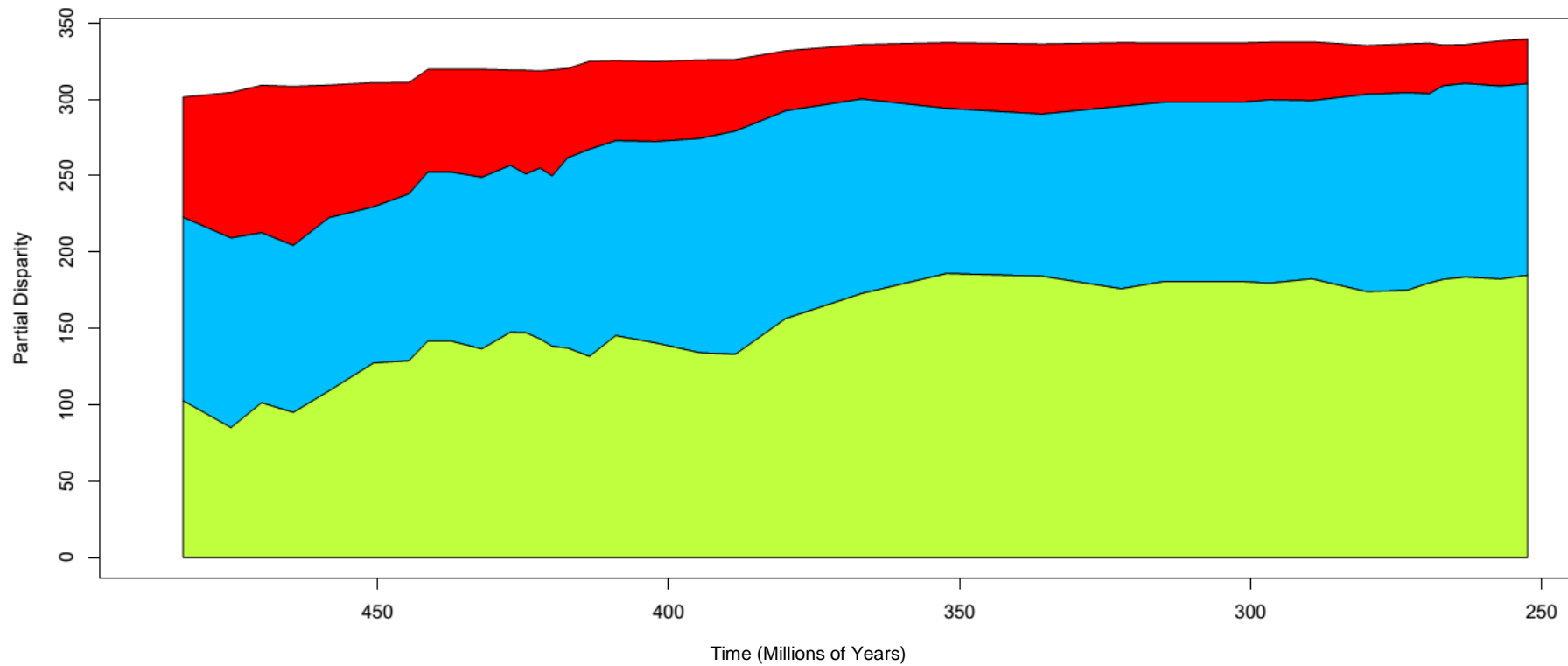


Figure.3.30: Partial disparity of subclades across the Palaeozoic. The top surface indicates the total level of disparity (sum of variances) of the three subclades combined across each stage interval Palaeozoic. The Mixed Range treatment was used to create this plot. The contribution of each subclades to the total disparity is represented by its colour. Red: Protobranchia; blue: Heteroconchia; and green: Pteriomorpha.

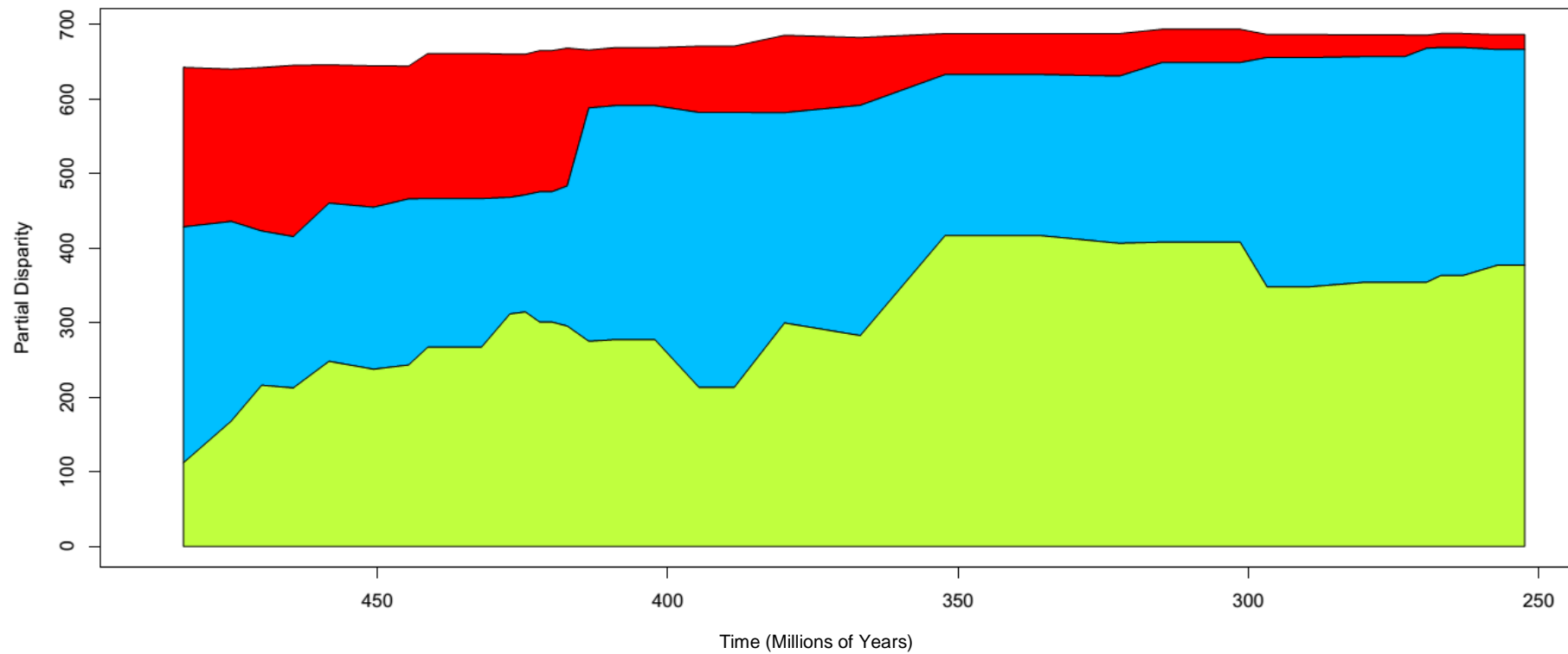


Figure.3.31: Partial disparity of subclades across the Palaeozoic. The top surface indicates the total level of disparity (sum of variances) of the three subclades combined across each stage interval Palaeozoic. The Treatise treatment was used to create this plot. The contribution of each subclades to the total disparity is represented by its colour. Red: Protobranchia; blue: Heteroconchia; and green: Pteriomorphia.

contribution made by Pteriomorphia in the Devonian and Permian, accompanied by an increase by the Heteroconchia (Fig.3.31). At the beginning of the Ordovician, each subclade in the Mixed Range treatment contributes roughly similar levels to the total disparity at that time (Fig.3.30). In the Treatise treatment the Heteroconchia contribute the most to disparity (Fig.3.31) but quickly shrink by the Middle Ordovician (Fig.3.31). By the End of the Permian both treatments suggest that Pteriomorphia contribute the greatest amount to the total disparity pattern, followed by the Heteroconchia and then the Protobranchia which contributed the least.

Comparing the partial disparity patterns of each subclade with its equivalent pattern produced in isolation (Fig.3.27-3.29) highlights some discrepancies between partial disparity and individual disparity patterns. This is most notable in the Protobranchia. The isolated pattern (Fig.3.27) shows disparity in the Protobranchia increases up to the end of the Silurian, stabilising thereafter. The partial disparity trajectory (Mixed Range and Treatise treatments) of Protobranchia suggests initially high disparity in the Ordovician – Silurian followed by either a gradual (Fig.3.30) or rapid (Fig.3.31) decline in the Late Silurian - Early Devonian and gradual decline up to the end of the Permian. The isolated disparity (Fig.3.28) and partial disparity (Fig.3.30 and 3.31) patterns of the Heteroconchia give seemingly contrary trajectories of disparity in the Earliest Ordovician.

The isolated disparity pattern shows an increase (Fig.3.28), the Mixed Range partial disparity pattern shows no change (Fig.3.30) and the Treatise partial disparity pattern shows a decline (Fig.3.31). The increase shown in the isolated disparity pattern in the Earliest Ordovician can coexist with the stable and decreasing partial disparity patterns (Fig.3.30-3.31) as the former isolated pattern provided the account of changing disparity within the Heteroconchia, while the latter patterns provide an account of how the disparity in the Heteroconchia was evolving relative to the rest of the Bivalvia. The isolated disparity patterns for the Heteroconchia (Fig.3.28) do not show any significant fluctuations across the Devonian or Permian as found in the Treatise (and lesser extent Mixed Range) partial disparity plots (Fig.3.30-3.31). This suggests the increase in the partial disparity is caused by a relative drop in the Protobranchia and Pteriomorphia rather than an increase in disparity within the Heteroconchia. For the rest of Palaeozoic the isolated and partial disparity results of the Heteroconchia agree that Heteroconchia disparity remains stable (Fig.3.27, 3.29-3.30). The morphospace plots of Figures 3.13 and 3.14 also back up these patterns, showing a lack of movement of the subclade through the bivalve morphospace.

There is also general agreement on the patterns of disparity (Fig.3.29) and partial disparity (Fig.3.30 and 3.31) for Pteriomorphia across the Palaeozoic. Both show that disparity is increasing in Pteriomorphia across the Palaeozoic, tapering off towards the Carboniferous-Permian (359.5-251 million years ago). The partial disparity plots pick up a short lived decline in the Early-Middle Devonian not immediately obvious in the disparity

through time plot (Fig.3.29). There is an increase in disparity in the Protobranchia during this interval (Fig.3.27). It seems that during the Devonian, there was a more complex interaction between the three subclades. Heteroconchia did not change significantly in individual disparity, but was contributing more to the total disparity pattern (Fig.3.30-3.31). At the same time the disparity and partial disparity within Protobranchia increased (Fig.3.27, 3.30-3.31), and the partial disparity in Pteriomorphia decreased.

### 3.28 Discussion

The results in this section provide a general account of the disparity of bivalves over the entire Palaeozoic. The disparity of bivalves grew across the Palaeozoic, increasing with minor perturbations caused by the end Ordovician and Late Devonian mass extinctions. The trend of increasing disparity was punctuated by sudden increases in disparity over time according to the results of marginal likelihood ratio tests (Tables 3.3-3.5) (Friedman 2010).. The exception to this general trend was found when comparing the disparity of survivors vs. extinction terminated taxa at the mass extinction boundaries. Marginal likelihood ratio tests of these taxa did find a significant increase in the disparity across those boundaries (Royall 1997; Finarelli & Flynn 2007; Friedman 2010). This result is similar to the results in other disparity studies that explore the effects of mass extinctions on disparity patterns (Friedman 2010; Kolbe *et al.* 2011) , i.e. that mass extinctions could potentially increase disparity by removing less disparate genera (Table 3.6).

Foote (1991a) suggest that this could be the consequence of a random extinction; however, in this case (Fig.3.14D and 3.15D) the increase in disparity over the Ordovician likely the result of a mixture of processes. The Pteriomorphia lost a number of genera in the middle of their occupied space whilst simultaneously gaining new forms at the periphery of their space furthest from the point of pteriomorphian origin (Fig.3.14D and 3.15D). During this time Heteroconchia drop in diversity, with a number of genera seemingly going extinct in a randomly distributed manner (Fig.3.14D and 3.15D). The Late Devonian mass extinction shows a similar pattern in the Mixed Range treatment (Fig.3.15E) whilst the Treatise treatment suggests this increase in disparity across the extinction was caused by the removal of two genera in the Pteriomorphia belonging to the Pterineidae: *Cornellites* and *Glyptodesma* (Carter *et al.* 2000; Carter *et al.* 2006; Carter *et al.* 2011). These two genera were found in near the origin point of the Pteriomorphia space. These results show that random or selective extinction selection can give the same result (increasing disparity). Therefore the effect mass extinctions have on disparity patterns is more complex than originally considered (Foote 1991a; Friedman 2010; Kolbe *et al.* 2011). The results describing the distribution of genera through morphospace (Fig.3.14-Fig.3.15, Appendix.3) per time intervals (Stage) show that Heteroconchia do suffer an extinction just prior (Givetian - Frasnian) to the Late Devonian extinction



boundary (Frasnian - Famennian). The distribution of those heteroconch genera that become extinct at this event falls in line with the random extinction selection scenario of Foote (1991a).

The morphospace of the Palaeozoic was highly clustered as shown by nearest neighbour and principal points analysis (Fig.3.22-3.25). This clustering has itself changed with time as genera have explored new portions of the bivalve morphospace (Fig.3.12-3.19). As disparity increased in bivalves as a whole, the disparity of subclades contributed to this increase to in a dynamic fashion. Protobranchia and Pteriomorphia increased in disparity when considered in isolation (Fig.3.27 and 3.29) while Heteroconchia remained stable (Fig.3.28). When the subclades were considered in tandem (Fig.3.30-3.31), they presented a more complex interaction of increases and decreases in partial disparity contribution across time.

The results of the movement of the subclades through the morphospace (Fig.3.14-3.19) and level of clustering across time (Fig.3.20-3.21) showed that all three subclades shared the same origin area of the morphospace, but branched out in different directions. The Pteriomorphia spread the furthest from the initial origin point but never fully vacated it. Instead the subclade spread outwards and became more clustered as they went, leaving areas of the space initially occupied empty or only containing one or two genera. The Heteroconchia reach the limits of their space by the end of the Silurian, and all change remains within the confines of that space (Fig.3.14-3.19). The Protobranchia began in two distinct clusters, one at the same origin point in the morphospace where Heteroconchia and Pteriomorphia appeared, the other away from either the Heteroconchia or Pteriomorphia (Fig.3.14-3.19). This likely contributed to their high level of partial disparity in the Ordovician (Fig.3.30-3.31) and relative decline as the Heteroconchia and Pteriomorphia explore more of the morphospace (Fig.3.14-3.19).

The disparity of Bivalvia through the Palaeozoic follows a distinctly different trajectory to the diversity of the bivalves as recorded by the traditional counts (Fig.3.8, 3.9A and 3.10A) or SQS (Fig.3.11A). These diversity curves show a peak in the Silurian and Permian respectively with a decline across the Devonian. The effect of the Devonian extinction on bivalve disparity seems to have been an increase in clustering in the Protobranchia (Fig.3.20BD, 3.21AD, 3.22AD) and Heteroconchia (Fig.3.21BE, 3.22BE). During this Period the sea level dropped (Fig.2.17) (Haq & Schutter 2002; Snedden & Liu 2010) in the first half, but rose again for a brief time in the Middle-Late Devonian, after which it went into major decline up to the Late Carboniferous where it was almost the same as present day sea level (Fig.2.17).

Through the Ordovician a number of feeding guilds evolved in the bivalves (Sanchez 2008). These guilds were all of endofaunal (shallow infaunal) life habits which fits with the data seen in Figure 2.17 of Chapter 2 (a high proportion of infaunal genera across the Phanerozoic); however, the study of Sanchez (2008) mainly looked at

Heteroconchia and to a lesser degree Pteriomorphia. Palaeotaxodonta ( $\approx$ Protobranchia) were part of the data but restricted to two families, one in each of the two basins of interest in the study (Sanchez 2008). The exploration of new feeding guilds by bivalves could be one factor in causing the rise in disparity of bivalves in Ordovician (Fig.3.26-3.29). The End Ordovician mass extinction does not seem to have any major effect on the isolated disparity of bivalves, except for a small drop in the Protobranchia (Fig.3.26); however, prior to this drop Protobranchia shared the greatest portion of the bivalve disparity (Fig.3.30-3.31).

They also occupied two distinct regions of space (Fig.3.14C-3.15C), which also show an increase in clustering (Fig.3.20B). One explanation could be that there are a number of Protobranchia forms missing from the data which would fall within the gap; however, as the sampling regime for the data tried to sample across the wide variety of Palaeozoic bivalves, leaving out poor quality data, I believe that the addition of other protobranchian taxa would not contribute to the gaps in the space (especially if the new material was of low quality e.g. worn, broken, incomplete).

The change in disparity across the Devonian prior to the mass extinction event is interesting as during this time bivalves were in decline (Fig.3.8-3.11) but disparity in bivalves rose. Subclade disparity rose (Fig.3.27-3.29), but heteroconchian partial disparity increased compared to Protobranchia and Pteriomorphia. This period would represent a good opportunity for future work to explain the increase in disparity, what clades were becoming extinct and what caused of the drop in partial disparity and clustering in Pteriomorphia and increase in clustering within the Heteroconchia (Fig. 3.21-3.22).

The morphospace occupation in the Devonian shows that the movement of Pteriomorphia through morphospace was caused by forms appearing at the outer edge of the space furthest from the origin. Simultaneously genera within the space became extinct (Fig.3.14-3.15). The Protobranchia lost a few genera though the Devonian but did not change in their position of the space (Fig.3.14-3.15). The continental set up of the Devonian was of two large continents: Gondwana and Euramerica (Laurentia and Baltica). Between these continents a number of tectonic barriers opened and closed through the Middle and Late Devonian, resulting in the reduction of interbasinal geographic faunal exchange (Rode & Lieberman 2004). These barriers could help increase disparity and clustering by blocking off habitats and leading to increased morphological differentiation. During the Late Devonian extinction event itself there is an increase in the level of cosmopolitan genera (Rode & Lieberman 2004) but there does not seem to be any changes in disparity or partial disparity that can be linked to this.

The results of the evolutionary modelling are interesting as they favour the Ornstein-Uhlenbeck trajectory of evolution (Table 3.8), and either a GRW or URW model depending on the scale of time binning (Stage, Epoch or Period) or clade of interest (Bivalvia, Protobranchia, Heteroconchia or Pteriomorphia) if the OU model was not

available. The original use of these models was to model the evolutionary trajectory of particular morphological traits (Hunt 2006; Butler & King 2004) and so their meaning in relation to disparity, the property of clades, is more difficult to ascertain. Each of the models suggested a different type of trait evolutionary i.e. directional evolution for GRW, genetic drift (URW), stasis (STA) and adaptive evolution (OU). The disparity of the clade is built upon a mosaic of these processes and, like with evolution of genes or individual characters (Sites Jr. *et al.* 1996), some of the mosaic will be congruent, while others will not (Dowton & Austin 2002). For the analyses of this Chapter an important assumption was made, that the disparity trajectories following any of the model (GRW, URW, STA or OU) were built upon multiple, and preferably the majority of, traits across the genera evolved either in a congruent and/or additive fashion toward the specific model.

What do these models mean at the clade level? Stasis is the simplest to visualise and translate up to the level of a clade, i.e., that disparity remained unchanged through the evolutionary history of the clade or put another way, reached its peak in disparity early in its evolutionary history. This peak may be the result of some constraint on form e.g. environmental or habitat specialisation, inability to invade new areas where diversification would otherwise be possible, or suppression by the competition from other organisms. The GRW model, directional change in disparity, i.e. net increase (or decrease) in disparity over the clades history may refer to the gradual release or decrease of new forms as constraints are removed. Given the scale, the URW probably does not reflect genetic drift alone but instead could be the result of repeated radiations into similar but less stable environments that become depopulated due to severe changes in the environment, e.g., tectonic activity opening/closing areas then subjected to re-colonisation.

Finally the OU model, adaption toward a selected optimum, at the level of disparity change in a clade likely has a different definition for adaptive optimums at the level of traits (Hunt *et al.* 2008; Butler & King 2004). It is unlikely that the evolutionary model represent any one force or process acting on the clade, as the OU model represents movement through the adaptive landscape (Butler & King 2004). At the level of clades this could represent the ability of the clade itself to adapt across the range of new environments opened up by the movement of the continents. Bivalves are first found in Gondwanan sediments of the southern hemisphere (Sanchez 2008), and have radiated outward in the Cambrian. With the movement of the continents several barriers to dispersal have come and gone (Rode & Lieberman 2004).

In bivalves disparity does not seem to decrease with extinction (Fig.3.26), therefore as these continental barriers are realised (Rode & Lieberman 2004) then bivalves are available to radiate across a range of new environments whilst maintaining the forms they have previously explored. As time progresses the number of major environmental barriers could drop, representing a slowing of the rate of disparity increase.

There are other possibilities to what the optimum means on at the level of disparity in clades, it could be the saturation forms that are mechanically feasible causing the levelling off by the end of the Permian, or could instead be the a constraint of the discrete character data, i.e. that differences between taxa can only grow so much before even newly evolved forms are no more different than those that preceded them. Finally the constraint resulting in the levelling off could be due to the evolutionary cost of new forms increases as the clade gets older, and so new genera converge on previously used forms or combination of forms.

There is the question of the applicability of these models to disparity; currently the AD parameterisation assumes a normal distribution in the possible step changes between the ancestor and descendent values (disparity value of one stage and the next) while the joint parameterisation assumes a multivariate normal distributions of step changes across the time series (Hunt 2006). A Shapiro-Wilk normality test shows that the trajectory of disparity in bivalves (and subclades) does not follow a normal distribution ( $p\text{-value} = 0.000198$ ); however, this may not pose a serious problem to the evolutionary modelling method if deviation from normality is low (Hubbard 1978) and given that the assumption of normality is assumed for the step change distribution rather than the time series itself (Hunt 2006). The deviations away from normality come from late in the time series (Carboniferous-Permian) where there is a drop in the rate of net increase in disparity.

The results themselves (Table 3.8-3.10) show some variation in the pick of model, which could be explained in part by the assumption of normality in the step change modelling but even so there seems to be some interesting trends picked out. Table 3.8 indicates that the disparity in bivalves and subclades follow the Ornstein-Uhlenbeck process i.e. a curve of initially rapid increase which tapers off. This trajectory fits with the results of Chapter 1 that suggest that bivalves follow a top-heavy disparity trajectory. The reason for the switch to the unbiased random walk model as the time interval binning increases likely stems from low resolution of changes in disparity as the interval size increased. As the time bins increased the differences between disparities increased. The modelling fitting suggests that the disparity increase followed more closely with an unbiased random walk than Ornstein-Uhlenbeck or generalised random walk at these higher binning intervals. The marginal likelihood results do not contradict the result as the trend can still be strong yet not the process or individual significant jumps.

### **3.29 Conclusions & Future Prospects**

The analyses here were able to give a general account of the disparity and diversity through the Palaeozoic. The general trend in both is one of increase. The disparity in bivalves is increased significantly by the Palaeozoic mass extinctions (End Ordovician and Late Devonian) but only when survivors vs. extinction terminated groups are analysed. When disparity between each interval is compared using all genera in the

intervals of comparison there are no significant jumps in disparity. This suggests that bivalves gradually explore their morphospace. The result of increased disparity across the Palaeozoic mass extinctions is interesting as they provide a new dynamic to the effects of mass extinctions on total biodiversity. The idea that mass extinctions can increase disparity in clades that survive is not new (Foote 1991a) but has only been demonstrated empirically in a limited number of studies using separate clades for separate extinction events of varying magnitudes (Friedman 2010; Kolbe *et al.* 2011). The analyses of this section are novel as they are the first to explore the effect of multiple mass extinctions of high magnitude on the same clade.

Below the surface of the main disparity pattern of the bivalves the individual subclades contribute to disparity dynamically across the periods. Protobranchia provides the greatest contribution in the Earliest Ordovician but Pteriomorphia have the biggest share by the end of the Palaeozoic. Heteroconchia contribute to a greater degree in the Devonian, but contribute a stable amount before and after. The evolutionary models used here suggest that disparity grew rapidly in the Early Palaeozoic, tapering off toward the End Palaeozoic. Although one could associate the chosen model with processes of adaptive radiation, the model is more likely summarising the exploration of the space by the individual subclades, i.e., Protobranchia rapidly exploring the edge of their space by the Early to Middle Ordovician, followed by the gradual exploration and occupation of the morphospace by the Pteriomorphia. The results here are a significant first step to a complete understanding of bivalves through the Palaeozoic but there is still much work to be done.

The same data used here to construct the bivalve morphospace is also being used to create revised phylogenetic trees of the relationships between the Palaeozoic bivalves (Carter *et al.* 2000; Carter *et al.* 2006) (personal communication, Dr Joseph Carter, 2013). Once published this revised relationship data can be used as a back bone with which to understand how disparity is changing across the tree of bivalves using other measures of disparity, for example: phylogenetic distances (Wagner 1997) or the morphological disparity index of Harmon *et al.* (2003). One area of future work could involve collection of morphometric data (landmarks, outlines) to trace changes in size disparity in Palaeozoic bivalves, an interesting area of research given the unique morphology found in fossil (Isozaki & Aljinović 2009) and modern giant bivalves compared to smaller bivalves.

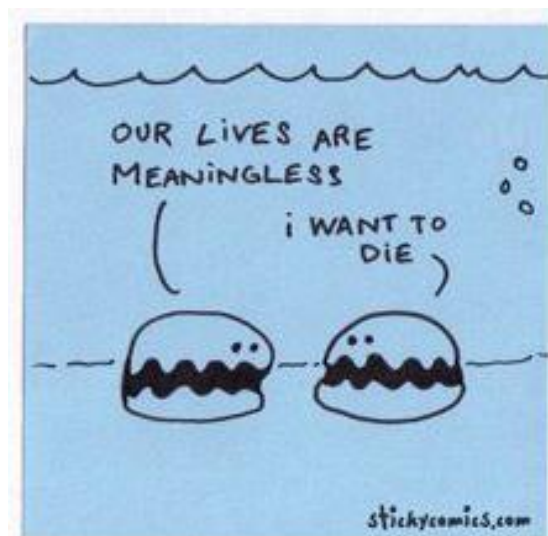
The disparity trajectories of this section covered the Palaeozoic; however, individual periods offer opportunities to more fully explore disparity across mass extinction events and between them. The evolutionary models tested the trajectory of disparity across the Palaeozoic; however, it is likely that the pattern is composed of numerous models each explaining different trajectories of individual traits in individual taxa. A sensible next step for any future projects would be to examine the evolutionary trajectories

of these traits to see how congruent they are, i.e. is the evolutionary model chosen for disparity the pattern of the majority of characters across the genera?

The disparity analyses here focused on the Palaeozoic, while other studies in the literature focused on the Mesozoic-Cenozoic (Grey *et al.* 2010; Kolbe *et al.* 2011). Given the semi-automation of the analyses developed as part of this thesis, future projects could expanded and revise the data to include bivalves from across the Phanerozoic. The only comparable dataset in the literature is the crinoid data of Foote (1999) which spans the Phanerozoic. Bivalves would make an excellent additional dataset with which to study disparity across the Phanerozoic when the morphological data becomes available. .

## ***Chapter 4***

### ***The Disparity of Bivalves Across the Globe***



#### 4.1 Introduction

Whilst the last chapter focused on changes in disparity and morphospace occupation of bivalves and constituent subclades through the Palaeozoic, this chapter looks at how disparity changes across latitude. Latitudinal diversity gradients are usually associated with changes from high density diversity in the tropics that drops towards the poles, but can refer to any changes in diversity with latitude. They have been observed and documented hundreds of times (Hillebrand 2004), and occur in numerous clades, from ocean plankton (Allen & Gillooly 2006) to crabs (Levinton & Mackie 2013), birds (Gaston 2000) and land plants (Gentry 1988). Latitudinal gradients can persist for millions of years. The diversity peak of the latitudinal diversity gradient for the modern marine biota is 30-45° North, and has persisted in this position since the Jurassic (Vinarskia *et al.* 2011).

Latitudinal diversity gradients are not static as shown in Figure 4.0. The Figure shows the palaeolatitudinal diversity gradient for the marine biota across the Phanerozoic (generated by the Paleobiology Database). The marine biota gradient can be tracked to the beginning of the Ordovician where peak diversity was between 35-40° South (Vinarskia *et al.* 2011). This peak shifted to 15-30° South until the end of the Devonian, and then gradually moved north again to its current position by the end of the Triassic (Crame 2002). The movement of this gradient with time can partly be explained by the movement of the continents, which begin mostly in the southern hemisphere (Fig.4.1AB) but gradually collide to form the supercontinent Pangaea by the end of the Permian (Fig.4.1C). Explanations for these gradients have been linked to changes in temperature, radiation, precipitation, potential evaporation, degree of vegetation, habitat, species ranges, and species incumbency (Levinton & Mackie 2013; Roy *et al.* 2000; Qian *et al.* 2009; Davies *et al.* 2007; Roy, Hunt, Jablonski, *et al.* 2009; Valentine *et al.* 2008; Markwick 2002).

Alone, these correlations cannot fully explain how latitudinal gradients in diversity are initiated and maintained over millions of years. It was not until the introduction of bivalves as a model organism that the evolutionary mechanisms which build and maintain these gradients were explored. The reasons for using bivalves are numerous, as previously discussed in Chapters 2 and 3. The latitudinal diversity peak in modern bivalves is at 25° North (Valentine & Jablonski 2010; Valentine *et al.* 2008; Roy *et al.* 2000; Krug *et al.* 2007), placing it firmly in the tropics (Fig.4.2). Other literature (Crame 2002) traces the bivalve latitudinal diversity gradient as far back as the Tithonian (Late Jurassic); however, it was far less steep than modern day. The rise between the Tithonian and modern day bivalve latitudinal diversity gradient was explained by the change from an equal number of epifaunal and infaunal taxa to increased levels of infaunal heteroconch taxa originating in the tropics (Crame 2002; Roy *et al.* 2004).



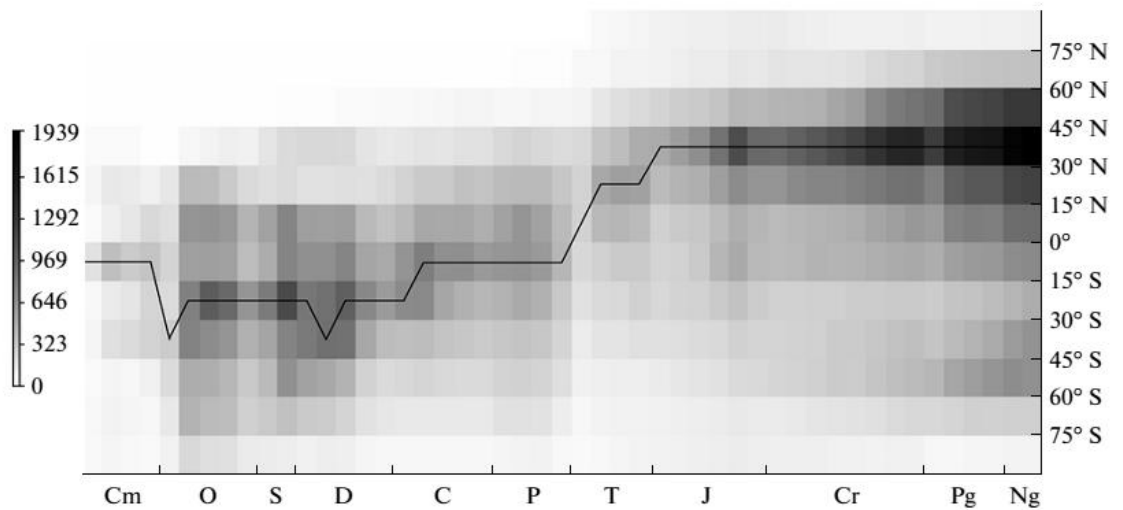
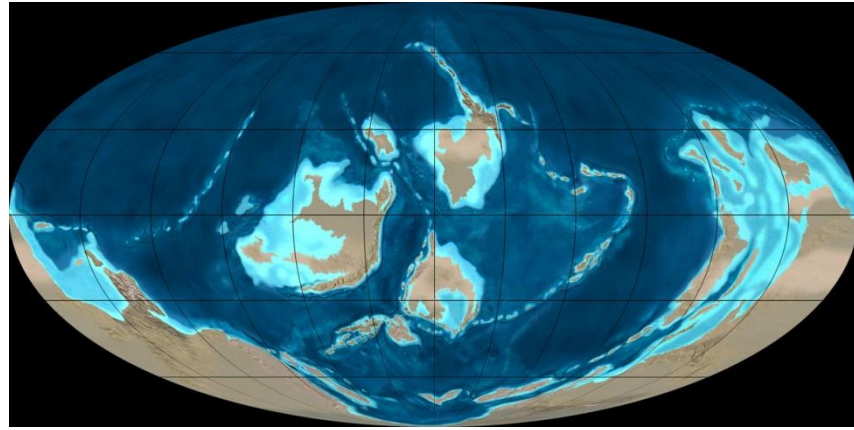


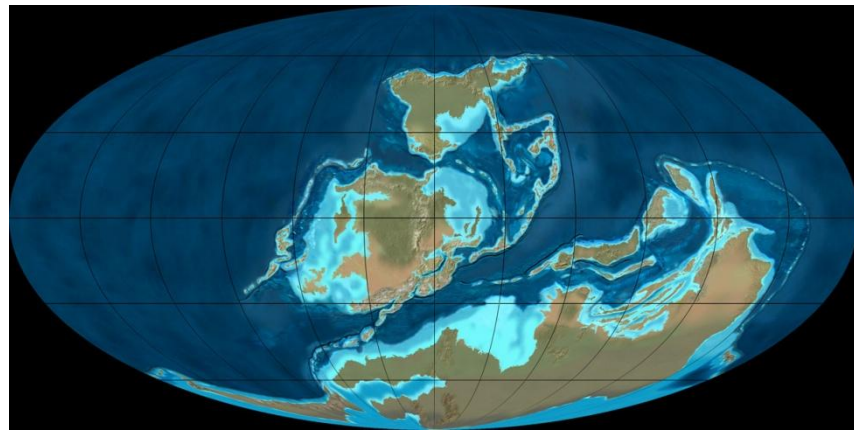
Figure.4.0 The latitudinal diversity gradient for the marine biota across the Phanerozoic. Latitude is split into 15°bins from southern to northern hemisphere. The black line shows the trajectory of the peak diversity across latitude in each time period. Time is measured in Periods, Cm: Cambrian, O: Ordovician, S: Silurian, D: Devonian, C: Carboniferous, P: Permian, T: Triassic, J: Jurassic, Cr: Cretaceous, Pg: Palaeogene, Ng: Neogene. Coloured squares of increasing intensity represent density of genera. Modified from Vinarskia *et al.* (2011).

Numerous studies have used the bivalve fossil record to test and/or affirm the tropics as a macroevolutionary source of diversity, i.e. the out of the tropics hypothesis: taxa originate faster in tropical latitudes than extra-tropical latitudes, migrating poleward but maintaining a link with the tropics (Vinarskia *et al.* 2011; Valentine & Jablonski 2010; Jablonski *et al.* 2006; Valentine *et al.* 2008; Krug *et al.* 2009; Roy, Hunt, Jablonski, *et al.* 2009; Roy & Goldberg 2007; Kiessling *et al.* 2010; Valentine *et al.* 2013; Mittelbach *et al.* 2007). They have also been used to show that extra-tropical latitudes can be a buffer against mass extinction (Vilhena *et al.* 2013).

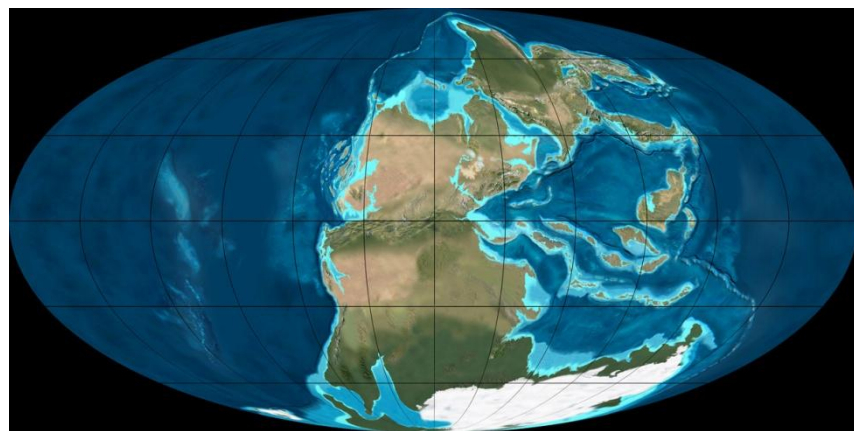
Despite the importance of latitude on taxonomic diversity, there is surprisingly little research on the changes in disparity with latitude. The reason for this is simple, few morphological datasets have been compiled that contain the necessary geographical (or palaeogeographical) data for assessing the changes of disparity with latitude. The first attempt to assess the disparity of a clade in relation to latitude began with a study of extant strombid (family Strombidae) gastropods (Roy *et al.* 2001). The study used shell shape and size as morphological variables, analysed via elliptical Fourier analysis. Combined with a compilation of geographical ranges taken from published maps the authors were able to quantify disparity per latitudinal bin (20° bins range from 40°South to 40°North). The authors found that morphological disparity fitted poorly with spatial patterns of species richness, finding that areas of high or low diversity could have high levels of disparity.



(A) Ordovician



(B) Devonian



(C) Permian

Figure.4.1 Mollweide globes showing the position of the palaeocontinents during the: (A) Ordovician, (B) Devonian and (C) Permian periods. Lines represent latitude and longitude of 30°bins. Sourced from Blakey (2011).

Neige (2003) focused on extant species of cuttlefish. Using morphological landmarks he quantified changes in cuttlebone shape, measuring disparity using the sum of ranges (Foote 1993b). Biogeographic units from across the globe gave no evidence for a gradient in diversity or disparity. Although the data was sampled geographically broadly, the emphasis of the study was on detecting gradients along the

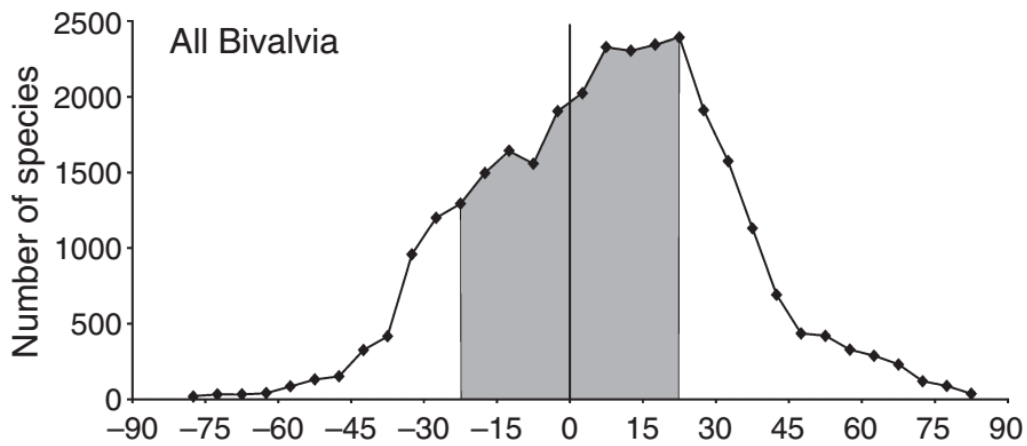


Figure.4.2 The latitudinal diversity gradient for modern bivalves. The gradient is based on counts of species in 15° latitudinal bins from the south pole to the north pole. Solid vertical line highlights the equator. The grey solid block highlights the tropics between 25°South to 25°North. Modified from Valentine *et al.* (2010).

sampling regime rather than specifically across latitude. Some latitudes were sampled more than others, and the latitudinal bins themselves were not studied in order. The studies of Roy *et al.* (2001) and Neige (2003) were good precursors for bridging the divide between disparity and latitude but shared similar caveats. Both datasets were restricted mainly to a thin band of the tropics (between 20°South and 20°North). Secondly, the emphasis of both studies lay with the longitudinal changes in disparity in this short band rather than latitudinal changes.

There have been two more recent studies that have made some effort in quantifying the disparity of clades relative to latitude. The first looked at the disparity of desmidiaceae (green algae) at a number of localities in the Czech Republic and Slovenia (Neustupa *et al.* 2009). Using outlines of the two planes of these algae, the study found a weak change with latitude. That is, higher latitudes contained greater partial disparity. Unfortunately this was never explicitly tested. The paper itself was always intended as a preliminary exploration of regional desmid disparity. The data was also restricted by latitudinal scope (45-50°North) and heavily biased sampling within the Czech Republic (Neustupa *et al.* 2009).

The other case study which looked at disparity across latitude was the study of Grey *et al.* (2010). This study, as discussed in Chapter 3, looked at the species level disparity of the genus *Buchia* through the Late Jurassic – Early Cretaceous (Oxfordian-Hauterivian). The dataset again found no evidence for a change in disparity with latitude. Although the conclusion was based on a large pool of measurements, these resolved changes in disparity within the genus only.

Fundamental questions remain pertaining to the interplay of disparity, latitude and time that have yet to be addressed. Does disparity follow a gradient similar to diversity? The current literature suggests that disparity does not follow diversity (Roy *et al.* 2001; Grey *et al.* 2010), yet there are issues with their findings. Roy *et al.* (2001) was mainly

concerned with longitudinal change and so only used data from between 20°North and South. Similarly Grey *et al.* (2010) relied on a single stage (the Tithonian) and as a result of sampling, was only able to determine disparity in the palaeolatitudes between 30-60°North and 30-60°South, with no data in between.

Bivalves have been used to show the tropics are a macroevolutionary source for diversity (Vinarskia *et al.* 2011; Valentine & Jablonski 2010; Jablonski *et al.* 2006; Valentine *et al.* 2008; Krug *et al.* 2009; Roy, Hunt, Jablonski, *et al.* 2009; Roy & Goldberg 2007; Kiessling *et al.* 2010; Valentine *et al.* 2013; Mittelbach *et al.* 2007), while the extra-tropics have been noted to act as a buffer to morphological extinction (Vilhena *et al.* 2013). If the tropics are a source of increased diversity which supplies the extra-tropics, this could increase disparity in tropical latitudes because strong competition for limited resources could lead to habitat specialisation (Schemske 2009). The extra-tropics may then have reduced disparity as taxa in these areas require more general diets, and so have similar morphologies to cope with the lower energy environments. Another possibility is that the extra-tropics could have higher disparity as clades in these areas become buffered from extinction events which hit more specialised tropical taxa (Vilhena *et al.* 2013).

Both the latitudinal gradients of the marine biota and bivalves have remained static from the Jurassic to present, but, prior to this, migrated from southern to northern tropical latitudes. Does disparity evolve across latitude with time, and in way similar to diversity? Published studies cannot comment on this issue they have looked at latitudinal trajectories of disparity at fixed points in time, either at the present (Roy *et al.* 2001; Neige 2003; Neustupa *et al.* 2009) or the past (Grey *et al.* 2010). The aim of this chapter is to address the questions of: (1) whether disparity follows a gradient similar to diversity, (2) to find the contribution that tropical and extra-tropical clades make to disparity, and (3) to test whether disparity evolves over time like diversity. The bivalve dataset of Chapters 3 makes an excellent case study as the data cover a significant portion of bivalve evolutionary history for which little is known regarding disparity. The stratigraphic and palaeolatitudinal data is also readily available.

## **4.2 Materials & Methods**

### **4.3 Data Collection & Disparity Calculation**

The morphological data used to construct disparity herein comes from the same source as Chapter 3, i.e., the discrete character morphological matrix of Carter (2000; 2006). As in Chapter 3, the data has been condensed to the genus level to take advantage of the generic level diversity and palaeolatitudinal data downloaded from the Paleobiology Database (Alroy *et al.* 2001; Kiessling *et al.* 2013). Unlike Chapter 3 which used two separate stratigraphic treatments (Mixed Range and Treatise), the analyses of this chapter will only use the Mixed Range data treatment. The Mixed Range treatment

contains the stratigraphic ranges of the Paleobiology Database, with updated ranges using the Krug *et al.* (2007) database and other published literature (Johnston & Goodbody 1989; Cope 1996; Liljedahl 1990; Schneider 1999; Perry & Larsen 2004; Sartori & Harper 2009). The first reason for using only this single treatment comes from the results of Chapter 3 that found only minor differences between the stratigraphy of the Mixed Range and Treatise treatments that did not change the overall result in disparity. The other reason for using the Mixed Range treatment was that the data of the treatment is essentially an updated version of the data contained in the Paleobiology Database and so meant little manipulation was required to make the data suitable for analysis.

The period of study is also the same as in Chapter 3. The first reason behind the choice was practical; the morphological data is biased toward this interval and therefore it makes sense to look at this interval specifically. The other reason was more exploratory; the Palaeozoic continental setup was very different from today (Fig.4.2) and the published literature shows the marine palaeolatitudinal diversity gradient begins in the southern hemisphere moving northward (Fig.4.0) in tandem with the movement of the continents. The palaeolatitudinal disparity and diversity gradients of bivalves during the Palaeozoic are unknown and so part of the analyses of this chapter test whether bivalve diversity follows the same trend as the marine biota (Fig.4.0) and whether the same is true of disparity.

The morphological and stratigraphic data were not split into the three subclades: Protobranchia, Heteroconchia and Pteriomorphia like the analyses of Chapter 3. The reason for not doing so was a practical one, as the data was partitioned into each Stage, and then along latitudinal bins, there was not enough data to also partition by each subclade and still have enough data to provide robust estimates of disparity. As in Chapters 1 and 3, disparity was calculated using the sums of variances (Foote 1991a) via the R script developed in Chapter 1 (Hughes *et al.* 2013). Time intervals corresponding to the stratigraphic Stage units from the *International Stratigraphic Chart* were used as the finest temporal resolution (Gradstein *et al.* 2004; Ogg *et al.* 2008). Partial disparity was also calculated using the method of Foote (Foote 1993a). The morphospace of bivalves were explored using the first 2 principal coordinate axis to show off the distribution of the genera in the space in relation to each latitudinal bin. As in Chapter 3, the occupation of genera in the morphospace was plotted using convex hulls.

#### **4.4 Latitudinal Partitioning**

Palaeolatitudinal coordinates for the genera contained within the morphological matrix of Palaeozoic Bivalvia (Chapter 3) were downloaded from the Paleobiology Database on the 28<sup>th</sup> of May, 2013 (Alroy *et al.* 2001; Kiessling *et al.* 2013). Using the Mixed Range treatment; for each Stage of the Palaeozoic the observed genera in that interval were isolated. The palaeolatitudinal coordinates were then isolated from the

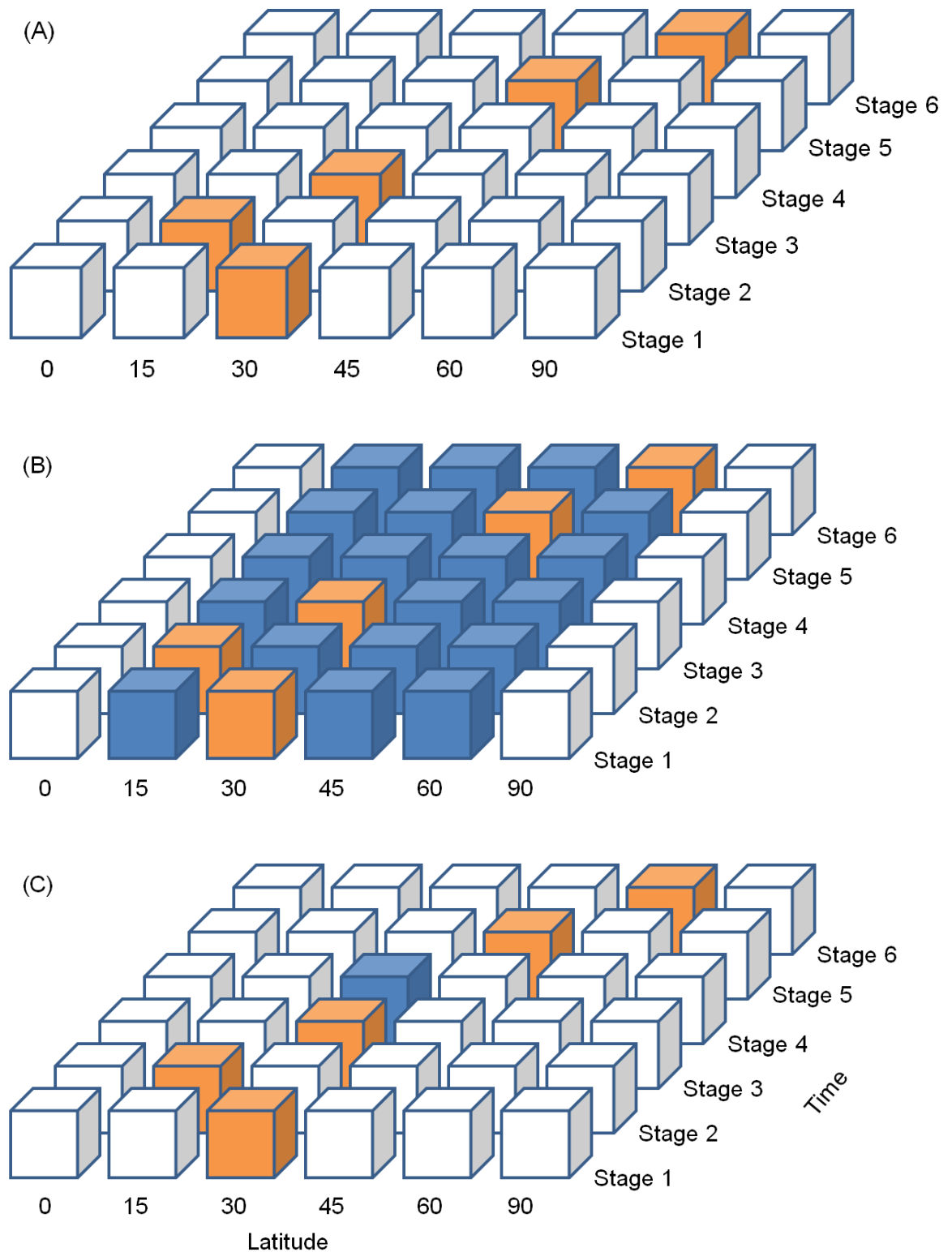


Figure.4.3 Hypothetical example showing the gap regime for the palaeolatitudinal data downloaded from the Paleobiology Database (northern hemisphere only). Each square is a possible presence of the genus as one latitude bin per Stage interval. (A) The range through time and space for one hypothetical genus. (B) Gap regime one assumes the presence of genus through its entire evolutionary history. (C) Gap regime two assumes that gaps in time are artifacts and fills in latitudinal range using data of the previous stage. Orange squares are the origin entry, blue squares are inferred entries based on the gap filling regime.

Paleobiology Database download. The Palaeozoic Earth was partitioned into twelve 15°bins ranging from the southern to northern poles (-90 to 90). The resolution was

chosen because latitudinal bins of this size (and smaller) have been used successfully in other latitudinal studies involving bivalves (Krug *et al.* 2009; Valentine & Jablonski 2010; Vinarskia *et al.* 2011). The 15°bin sizes were also considered the best compromise between providing the finest resolution of changes in disparity across latitude and time while still providing enough taxa in each bin to calculate disparity.

The latitudinal data was constructed into a presence (1) absence (0) matrix for each stage, split by the 12 latitudinal bins (Fig.4.3). The patchy nature of the fossil record (and the Paleobiology Database) meant that some latitudinal entries were missing for certain genera even though there was stratigraphic information available. These gaps were filled in an intelligent manner to provide a more robust dataset with which to assess changes in latitudinal diversity and disparity. Two different gap filling methods were used which made different assumptions about the quality of the fossil record. Both methods assumed that the palaeolatitudinal coordinates for each genus were incomplete (Fig.4.3A), but to different degrees (Fig.4.3BC). The first gap filling method (Fig.4.3B) assumed that the latitudinal coordinates of each genus were incomplete in time and in space and that the genus was present across all latitudes for the whole of its stratigraphy range. This method therefore assumed that the fossil record and/or Paleobiology Database heavily underestimated the palaeolatitudinal range of each genus.

The second gap filling method (Fig.4.3C) was more conservative, only filling in those gaps in palaeolatitude for stages between the first and last occurrence that were also missing latitudinal information. When these gaps were identified, the palaeolatitude for the Stage was assigned the value of the prior Stage. This second method relied on the assumption that the palaeolatitude records for each interval reflect the real latitudinal ranges of the genus when it was alive, and are not the product of a patchy fossil record. The two gap filling methods have their pros and cons. The first method ensures no missing palaeolatitude information, but could overestimate the palaeolatitudinal ranges of the genera across individual Stages. The second method also ensures no missing latitudinal information remains but may underestimate the degree of patchiness in the fossil record in individual Stages. Both methods also make the assumption that the extremes in latitude already found are the maximum and minimum limits of the palaeolatitudinal range of each genus, regardless of how patchy their fossil records are.

Once the palaeolatitudinal data had been filled for the genera in each Stage of the Palaeozoic, the distance matrix was produced using the R script of Chapter 1 and 3, the disparity calculated and the morphospace ordinate using principal coordinates analysis (Hughes *et al.* 2013). Diversity was also calculated using the traditional counts and the sample modelling proxy methodology (Lloyd 2011) of Chapter 2. To allow for enough data to calculate robust levels for the sampling proxy (number of formations and collections through time) this method was calculated using period level stratigraphy, i.e., finding the genera, collections and formations present in each period. The traditional count method

was used to look at changes in diversity per Stage. These two methods were chosen as they each convey different aspects of changes in diversity, i.e., the moving peak in diversity can be found using the traditional count method while the modelling method can determine the latitudinal bins in each period; that is, show diversity increases or decreases outside the amount expected as the result of sampling bias. The patterns in diversity found using shareholder quorum subsampling could be a potential area for future analyses as it may find peaks and troughs different to those found by the traditional counts method (Chapter 2).

#### **4.5 Tropical vs. Extra-tropical Clades**

The data was partitioned into three groups based on their palaeolatitudinal range to assess the levels of disparity contributed by each. The three groups included those genera that only occupied the tropical latitudes, those that only occupied the extra-tropical latitudes and those genera that ventured into both at least once along their stratigraphic range. The modern tropical range is between 23°27'South and 23°27'North. Due to the slightly wider ranging latitudinal bins of the analyses herein, the tropics were classified between 30°South to 30°North. The extra-tropical latitudes were classified between 30°-90°South and 30°-90°North.

Once the genera of the three groups were identified, the morphological data and stratigraphic ranges for those genera were subsampled from the Bivalvia morphological and stratigraphic data (Mixed Range) of Chapter 3. The distances were calculated using the same *R* script for Chapters 1 and 3, the morphospace was ordinated using principal coordinate analysis and the morphological disparity of the clade calculated using the sum of variances. The contribution of the three groups (tropical, extra-tropical and both) to the total disparity pattern was also analysed using partial disparity analysis (Foote 1993a).

#### **4.6 Results**

##### **4.7 Disparity Across Latitude – General Pattern**

Disparity was calculated across latitude over each stage of the Palaeozoic for bivalves (Fig.4.4-4.5). In Figure 4.4, the plot shows the interplay between disparity through time and across palaeolatitude based on the two gap filling regimes. Disparity in each stage of the Palaeozoic was calculated across the globe using 15° bins (12 in total) scale ranging from the southern (-90°) to northern poles (90°). The Cambrian was not included due a lack of stratigraphic resolution and low sample size per latitudinal bin, making latitudinal disparity calculation impossible. The overall trend in disparity using gap regime one and 2 (Fig.4.4) reflects the gradual increase in disparity with time that was a



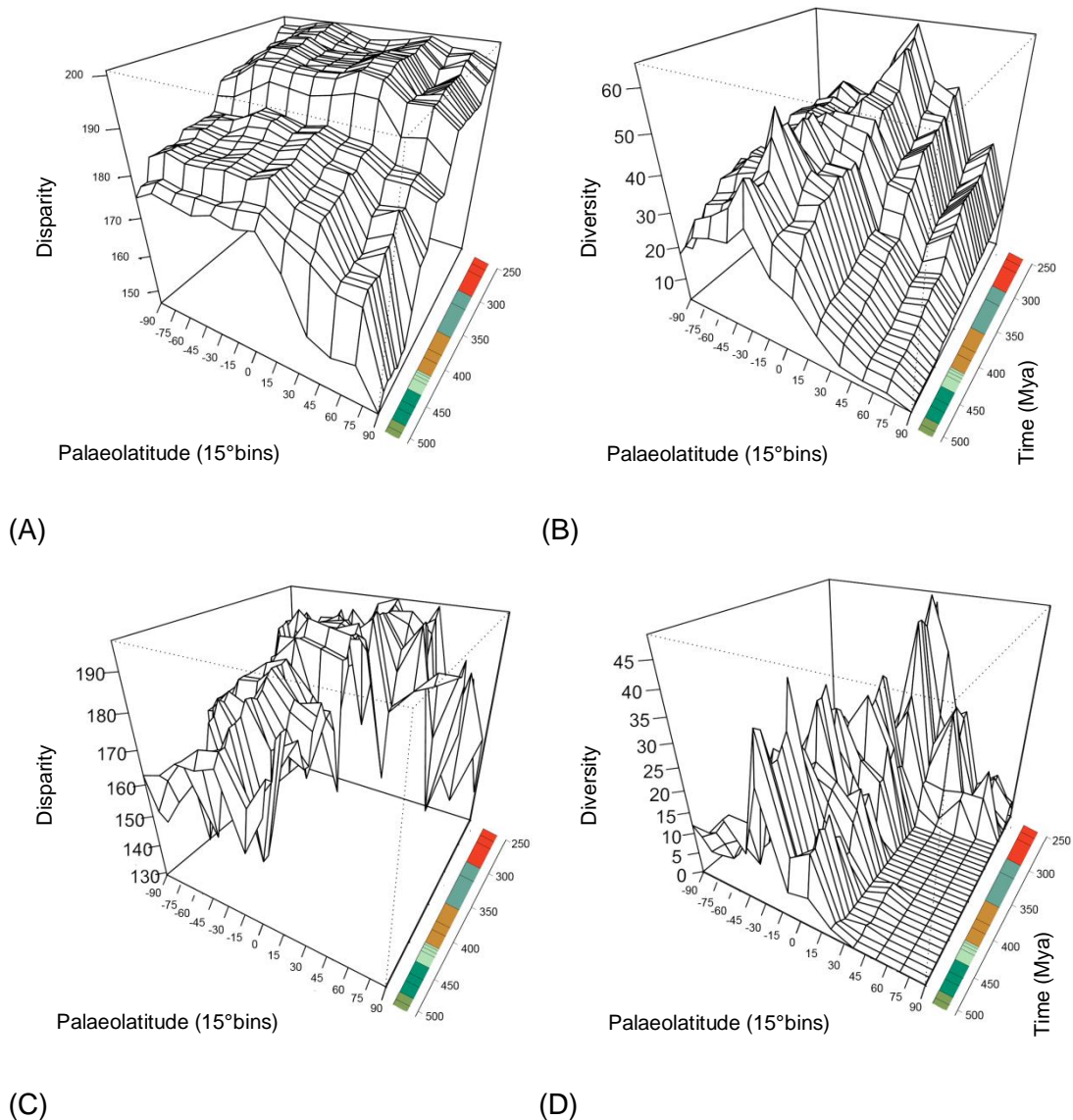


Figure.4.4. A three dimensional plot showing the change in disparity with time (Ordovician – End Permian) and palaeolatitude (15° bins from -90 to 90). (A) Disparity using latitude gap regime one. (B) Diversity (traditional counts) using latitude gap regime one. (C) Disparity using latitude gap regime two. (D) Diversity using latitude gap regime two.

prominent pattern in Chapter 3. Across latitude the trend in disparity was very different. In the early intervals of Fig.4.4A there was a drop in disparity from the equator up the northern latitudes while in the southern latitudes the disparity remained relatively stable. Due to low sample size, in Figure 4.4C (gap regime two) disparity could not be calculated at all for these northern latitudes.

From the Ordovician to the End of the Silurian the disparity in the northern latitudes rose (Fig.4.4A). At the same time disparity continued to remain unchanged. By the Devonian disparity was stable across all latitudes. For the rest of the Palaeozoic the global disparity grew but latitudinal disparity remained invariable across the latitudinal bins. The diversity across latitude and time (Fig.4.4B) presented a different pattern. From the beginning of the Ordovician diversity followed a gradient, peaking in the southern latitudes (45-30°South) and dropping away either side of the peak toward the poles. This

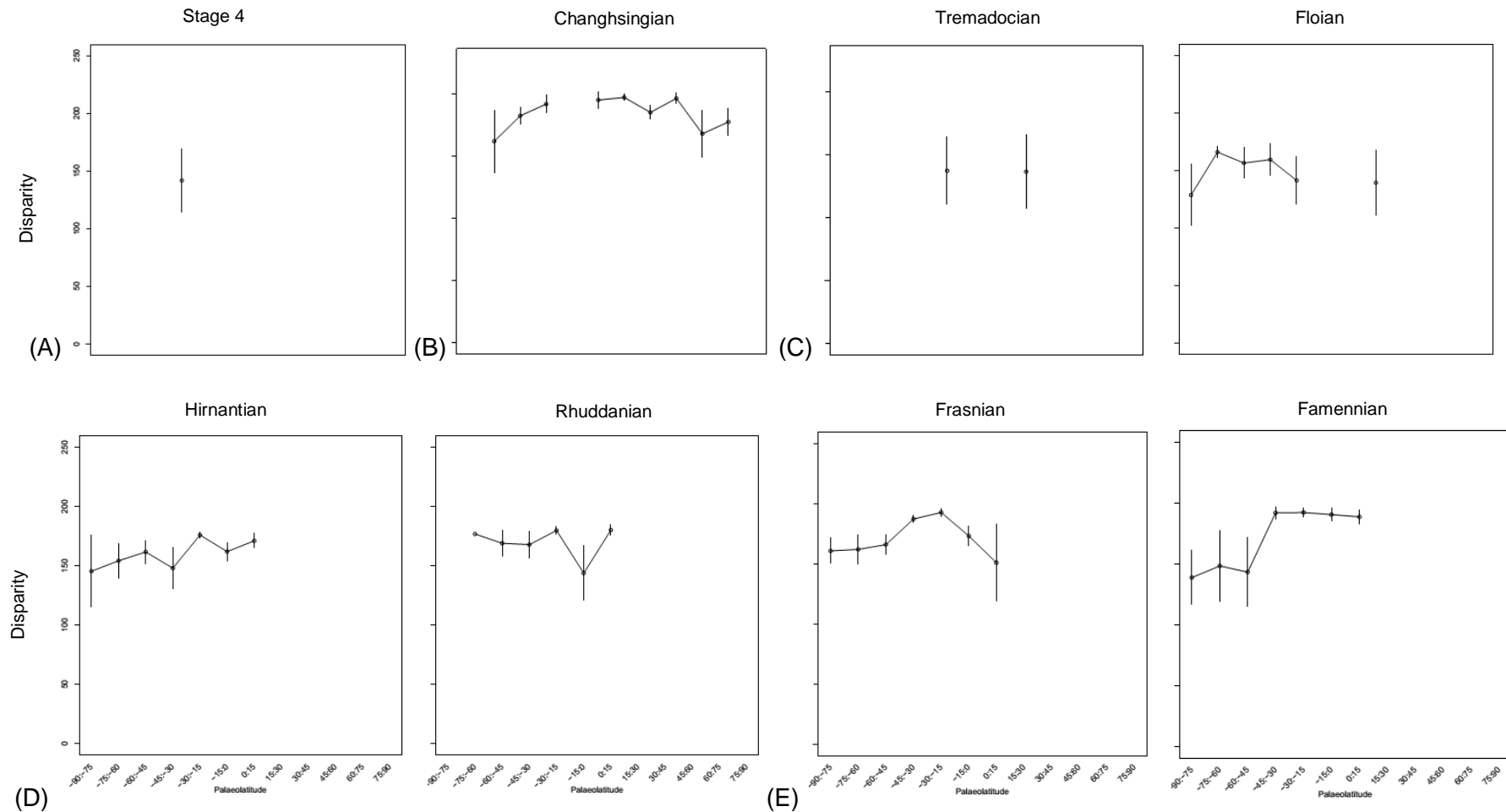


Figure.4.5. Disparity profiles for bivalves across palaeolatitude at specific stages of the Palaeozoic. The Mixed Range treatment was used. (A) Cambrian Stage 4 (~515-510 million years ago). (B) Changhsingian boundary (253.8-251million years ago). (C) Tremadocian - Floian: Early Ordovician (488.3-471.8 million years ago). (D) Hirnantian - Rhuddanian boundary (445.6-439million years ago). (E) Frasnian - Famennian boundary (385.3-359.2million years ago). Black dots are the mean disparity based on 1000 bootstrap replicates. Black error bars represent standard error. Blank parts of the curve are palaeolatitudes that do not contain the genera found in those stages. Palaeolatitude information downloaded from the PBDB.

gradient is comparable to the global palaeolatitudinal diversity gradient seen in the literature (Fig.4.0) (Vinarskia *et al.* 2011). As time passed the diversity peak gradually migrated from its southern latitude position to rest at the equator in the Permian. Through the Devonian and Carboniferous the diversity peak from previous stages dropped and widened from 1-2 latitudinal bins to 2-4 bins. This corresponds with the Devonian drop in diversity (Chapter 3 Fig.3.11A). The diversity results indicate that in bivalves, the Devonian decrease in diversity was not localised to any single latitude bin but instead decreased across numerous latitudes.

The disparity values using the second gap regime (Fig.4.3C) were sparser, with gaps of missing disparity values across the Palaeozoic. Where disparity could be calculated there were no differences in disparity values across the latitudes just as in the more complete data of Figure 4.3A. Figure 4.4 shows individual time slices Figure 4.3C. The disparity across latitude is more uneven as shown in Figure 4.3C due to the sparse data; however, where numbers of genera are abundant enough to give small standard error bars around disparity values, there was very little difference in disparity across latitude. The individual time slices for the first gap regime also showed unchanging disparity across latitudinal disparity through the Phanerozoic as in Figure 4.3A.

#### **4.8 Detrended Diversity**

The diversity of the bivalves, detrended via the sampling proxy modelling method of Lloyd *et al.* (2011; 2012; 2013) as described in Chapter 2 (see section **2.5 Quantifying diversity**) are shown in Figure 4.6. There are little differences between using either the number of collections or formations, so the number of formations is used to illustrate the differences between the two gap filling regimes. The method used the number of formations present in each latitudinal bin per Period as a proxy for sampling bias. For each Period, the method applied several regression models to the number of formations across latitude: linear, logarithmic, exponential, hyperbolic, sigmoidal and polynomial. Once applied the best fitting model was chosen using the Akaike Information Criterion (Johnson & Omland 2004) and expected diversity calculated. The expected diversity was then subtracted from the observed and all diversity outside of the error bars were taken as significant deviations requiring an explanation other than sampling bias.

The coarseness of gap regime one (Fig.4.6A) suggests in the Ordovician that there are more genera along the equator than expected by the sampling proxy and latitudes above 30°North (i.e. northern extra-tropical latitudes) the result of sampling bias alone. The high latitudes of the southern hemisphere (above 30°South) shows a lack of diversity given the sampling proxy. For gap regime two, there is little data for northern latitudes (Fig.4.6B). The amount of the data at different latitudes seems to fluctuate, with individual bins containing many more genera compared to the sampling proxy while others have far fewer.

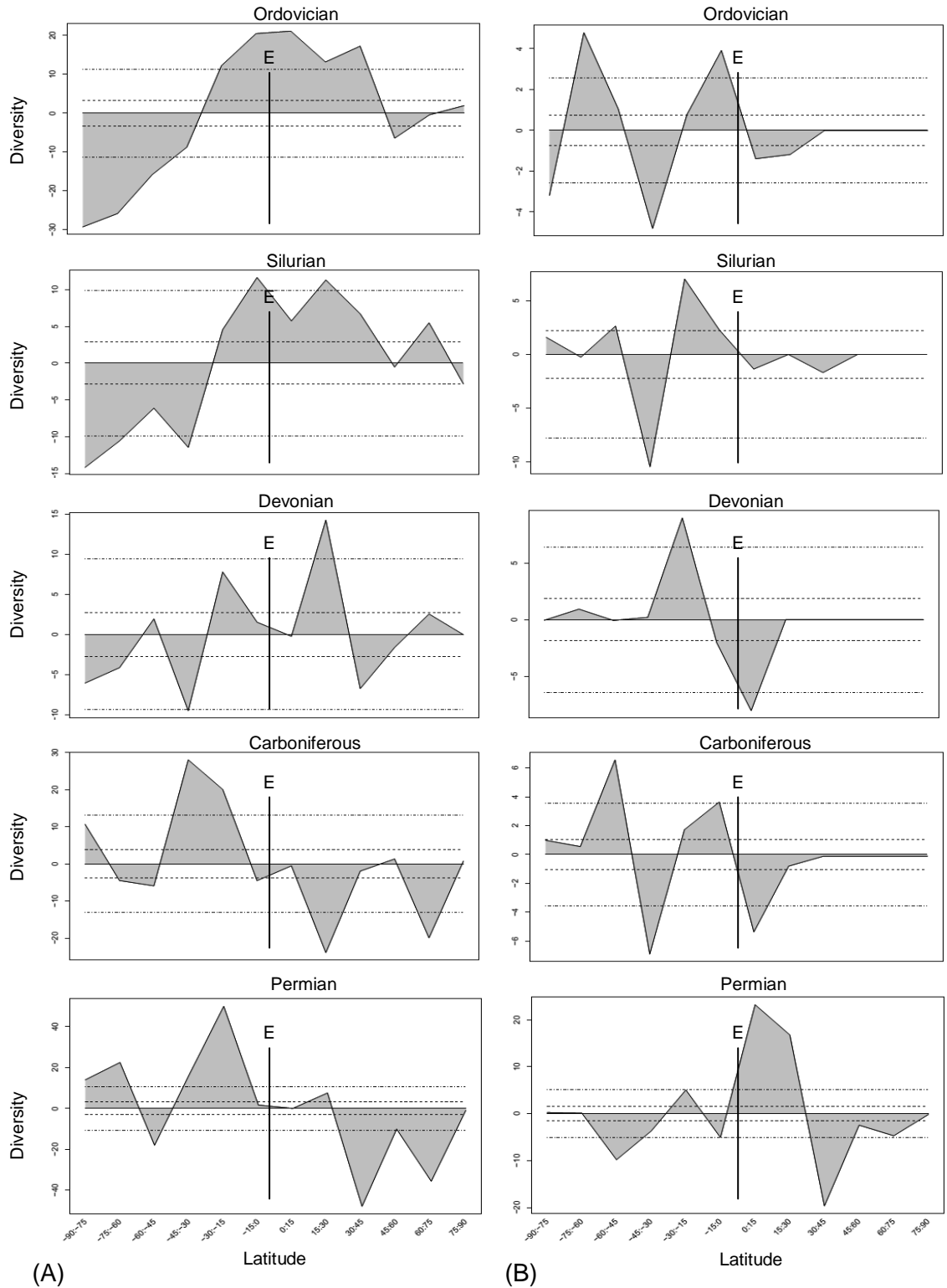


Figure 4.6. Detrended diversity across latitude for each period (Ordovician – Permian) of the Palaeozoic using the number of formations as the sampling proxy. Oldest period at top of page; youngest at bottom. The latitude is split into 15° bins from 90°South (left side of equator line) to 90°North (right side of equator line). The equator is highlighted with a solid line and the letter E. (A) Gap regime one. (B) Gap regime two. Horizontal dashed lines represent the 95% confidence limits (standard error) of diversity due to fossil sampling bias alone.

The Silurian (Fig.4.6) is much the same as the Ordovician; however, many of the latitudinal bins find the detrended diversity to be the result of the sampling proxy. The Devonian pattern also finds much of the latitudinal diversity of the Silurian using gap regime one to be the solely the result of the number of formations except for one bin in the northern hemisphere close to the equator. The Devonian results of gap regime two (Fig.4.6B) found a peak in diversity in the southern hemisphere close to the equator and a severe drop in diversity just north of the equator. The data using gap regime one in the Carboniferous found diversity rose above the sample proxies 95% confidence limit at the edge of the southern hemisphere tropic zone (defined as 30°South) (Fig.4.6A) while in the northern hemisphere the diversity of bivalves across latitude dropped within the range of fossil sampling bias and at 0-15°North dropped below the range, suggesting less diversity than expected. The result of the Carboniferous in Figure 4.6B suggest the diversity was higher than expected outside of the southern tropical latitudes (>30°South) but closer to the equator the diversity dropped into the range of fossil record sampling bias and into the area where diversity is much lower than expected. The Permian provides the most complete record across latitude as both Figure.4.6A and 4.6B can track diversity across the whole latitudinal range.

The result of Figure.4.6A for the Permian suggest diversity was far higher at the edge of the tropics in the southern hemisphere (30-15°South) than expected by the sampling proxy. The gap regime two edited data suggested that diversity north of the Equator (0-15°North) was far greater than caused by fossil record sampling bias; however, the data also suggest that at the edge of the northern tropics (30°North) the diversity dropped quickly below the level expected to be caused by sampling bias. The result of Figure 4.6 suggests overall that the quality of the taxonomic diversity record across latitude is variable, with some latitudes demonstrating higher than expected peaks in diversity and others low. One future refinement to the procedure would be to remove those formations that come from Lagerstättë. These types of formation could explain some of the individually high peaks in the results of Figure.4.6B. Removing these particular formations would have some affect on the results of Figure.4.6A; however, I suspect that due to the coarseness with which the first gap regime fills in missing latitude information, the impact of removing Lagerstättë on those results would be less severe. It would not affect the overall pattern unless those formations were situated on the outer edges of the latitudinal range of the genera.

#### ***4.9 Disparity Across Latitude – Tropic vs. Extra-tropical Disparity***

The general pattern in disparity across latitude in the Palaeozoic seems to be one of stability (Fig.4.4-4.5), except for the gradual increase with time that co-occurs across all latitudinal bins (Fig.4.4A, and Chapter 3 results). It has already been noted that underlying the general increase, subclade disparity varied across time, with taxonomic subclades

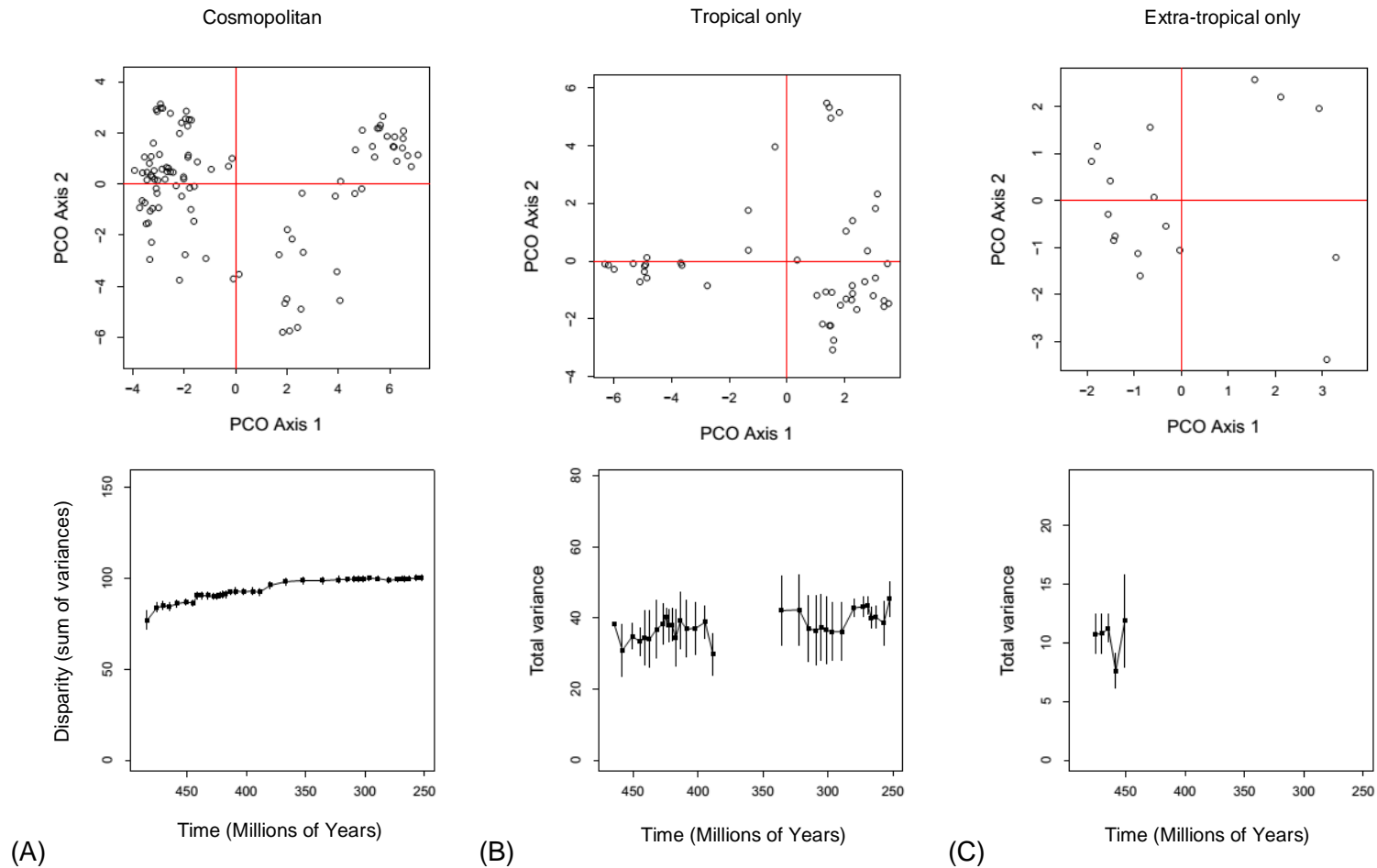


Figure.4.7: Palaeozoic disparity of bivalves split by latitude (Mixed Range treatment). The morphospace is conveyed by principle coordinate axes 1 and 2; each point is one genus of the subgroup. Disparity is measured using the sum of variances across the Palaeozoic; black: observed disparity, error bars based on 1000 bootstraps; green: observed disparity; red disparity based on random sampling from entire morphospace. (A) Cosmopolitan genera i.e. those genera that occur both in tropical AND extra-tropical latitudes (minus those clades that are found only in the tropics or extra-tropics). (B) Tropical latitude genera. (C) Extra-tropical latitude genera only.

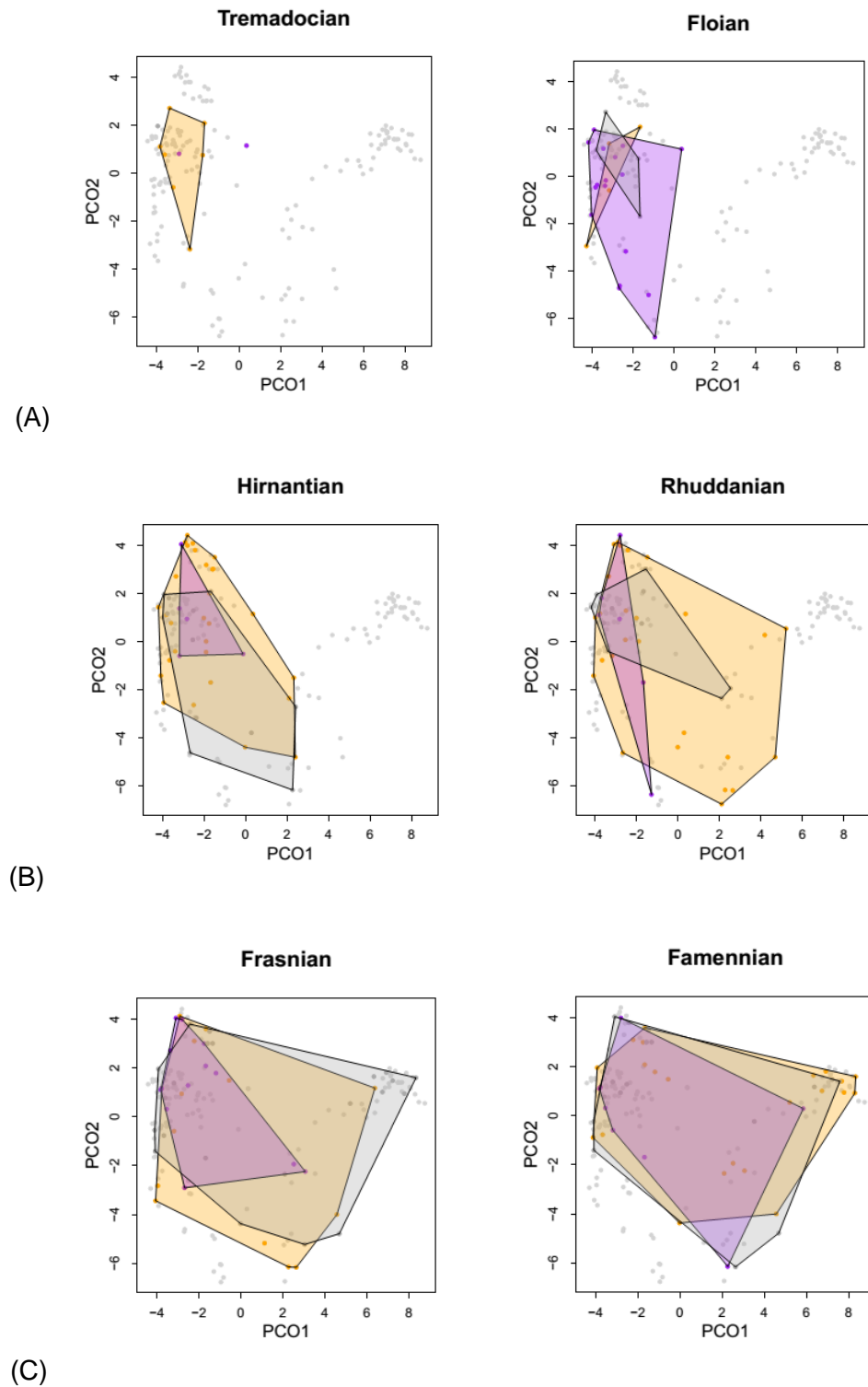


Figure.4.8. Morphospace occupation (convex hull) of genera included in the latitudinal subgroups across specific Stages of the Palaeozoic. (A) Early Ordovician Radiation (Tremadocian-Floian). (B) End Ordovician mass extinction (Hirnantian-Rhuddanian). (C) Late Devonian mass extinction (Frasnian-Famennian). Colours of convex hull represent the subgroups. Orange: tropical latitude clades only; purple: extra-tropical latitudes only; grey: cosmopolitan (appeared in tropics and extra-tropics).

contributing dynamically to the total disparity pattern (Chapter 3 Fig.3.29-3.30). The next set of results provide an account of disparity trajectories for those genera confined exclusively the tropics, extra-tropics and those wide ranging enough to be have occupied the tropics and extra-tropical latitudes at some point in their evolutionary history.

The data for the northern and southern hemispheres were combined. The tropics were defined as between the latitudes of 30°South to 30°North, while the extra-tropical latitudes contained all latitudinal bins from the range 30-90°South and 30-90°North. This was done to quantify trends caused by changes in latitude rather than specific differences between the northern and southern hemispheres. A future refinement in any follow-up studies could split the data of cosmopolitan genera into those from the northern and southern hemispheres, and so address any patterns exclusive to those hemispheres.

Given that the results of disparity across latitude using the two different gap regimes did not change the overall result of the disparity pattern across latitude (Fig.4.4), the rest of the results reported herein concern only the second gap regime data. The reason I chose this over the data modified by gap regime one was that the second regime only modified the latitudinal occurrences where necessary to fill missing entries within the stratigraphic range of the genera rather than draw out the genera into latitudes they may not have occupied across their entire stratigraphic range. The second regime essentially allows the genera themselves to have migrated across latitude across time.

The majority of genera fall into the cosmopolitan category (Fig.4.7A). The disparity trajectory of those genera is similar to previous analyses (Chapter 3 Fig.3.25, 3.29-3.30). The cosmopolitan genera do not cluster together in the morphospace. Instead they present a pattern similar to the morphospace of the Bivalvia (Chapter 3 Fig.3.12). The disparity of those exclusive to tropical latitudes is found in Figure 4.7B. The disparity trajectory shows no evidence for a trend through time. There was a gap in disparity through the Early Devonian to the Early Carboniferous. This drop is likely linked to the prolonged decline in bivalve diversity through this period (Chapter 3 Fig.3.11) resulting in a drop of genera, and therefore the ability to calculate disparity. The morphospace of tropical only genera shows a pattern like the one for cosmopolitan genera, simply mirrored. Those genera that were exclusively extra-tropical were the smallest group (Fig.4.7C). There is only enough generic data to calculate disparity in the Early Ordovician for this subgroup (Fig.4.7C). In this interval the disparity in extra-tropical bivalves did not change (large overlapping standard error bars). The morphospace for these extra-tropical genera suggest three clusters, probably corresponding to the three taxonomic clades.

The results of the subgroup disparity suggest that most bivalve genera were found across a wide range of latitudes. The way the analyses here collated cosmopolitan genera means that all those genera that had moved into new latitudinal bins and vacated others were binned into the same category as those genera that were truly cosmopolitan. Future refinements of the data could include identifying those genera that vacated their original



latitudinal position and those that continued to occupy them once established across both latitudinal expanses (tropics and extra-tropics). As the morphospaces of the isolated groups cannot be compared directly, the occupations of the subgroups were considered in tandem using the whole bivalve space (Fig.4.8). The Figure shows that through the Early Ordovician (Fig.4.8A), the tropical only clades surround the origin point of the space. From this point the clades expand, but this expansion is greater in the subtropical clades (larger convex hull) (Fig.4.8A), while those clades found in both tropics and extra-tropics or just the tropical latitudes remain near the origin of the space.

The end of the Ordovician (Fig.4.8B) involved an expansion in the morphospace occupied by tropical only genera, while there was a drop in the number of cosmopolitan genera. The expanse of the extra-tropical genera also increased across the End Ordovician mass extinction boundary. (Fig.4.8B). Across the Late Devonian mass extinction boundary there was an increase in the morphospace occupied by extra-tropical genera while there was little change in the tropical only and cosmopolitan genera (Fig.4.8C). The results show that all three subgroups overlapped in morphospace (PCO axes 1-2), and likely constitute genera from all three subclades of Bivalvia (Chapter 3 Fig.3.12).

As well as calculating the disparity of latitudinal groups in isolation, partial disparity analyses of the three subgroups are found in Figure 4.9. The partial disparity pattern is interesting as it shows a complex latitudinal interaction in disparity leading to the overall rise in disparity. The tropical only genera contribute the greatest to the disparity of the bivalves in the Earliest Ordovician (Tremadocian) (Fig.4.9). This is seen in the morphospace of Figure 4.8A where tropical only genera make up the full extent of the morphospace during that time. There was a switch in the Floian where extra-tropical genera contributed the most (Fig.4.9 and Fig.4.9A). Cosmopolitan genera began to increase in their contribution along with those tropical only clades through the rest of the Ordovician. This extended into the Middle Silurian (Fig.4.9). There was an increase in the contribution of tropical only clades across the End Ordovician mass extinction, also seen as an expansion of the morphospace occupied by the tropical only genera (Fig.4.8B).

In the late Silurian the contributions from extra-tropical genera increased (Fig.4.9) with a simultaneous drop in cosmopolitan and tropical only genera. Through the Early and Middle Devonian all three subgroups contributed the same amount to disparity but in the Late Devonian the disparity contributed by cosmopolitan genera decreased while tropical only genera increased. Nevertheless, during this time the degree of morphospace occupation of cosmopolitan genera did not decrease, instead the tropical genera rotated and grew through the morphospace (Fig.4.8C). Through the Early Carboniferous there was an increased contribution by the tropical only genera (Fig.4.9) with a drop in the contribution made by the cosmopolitan genera.

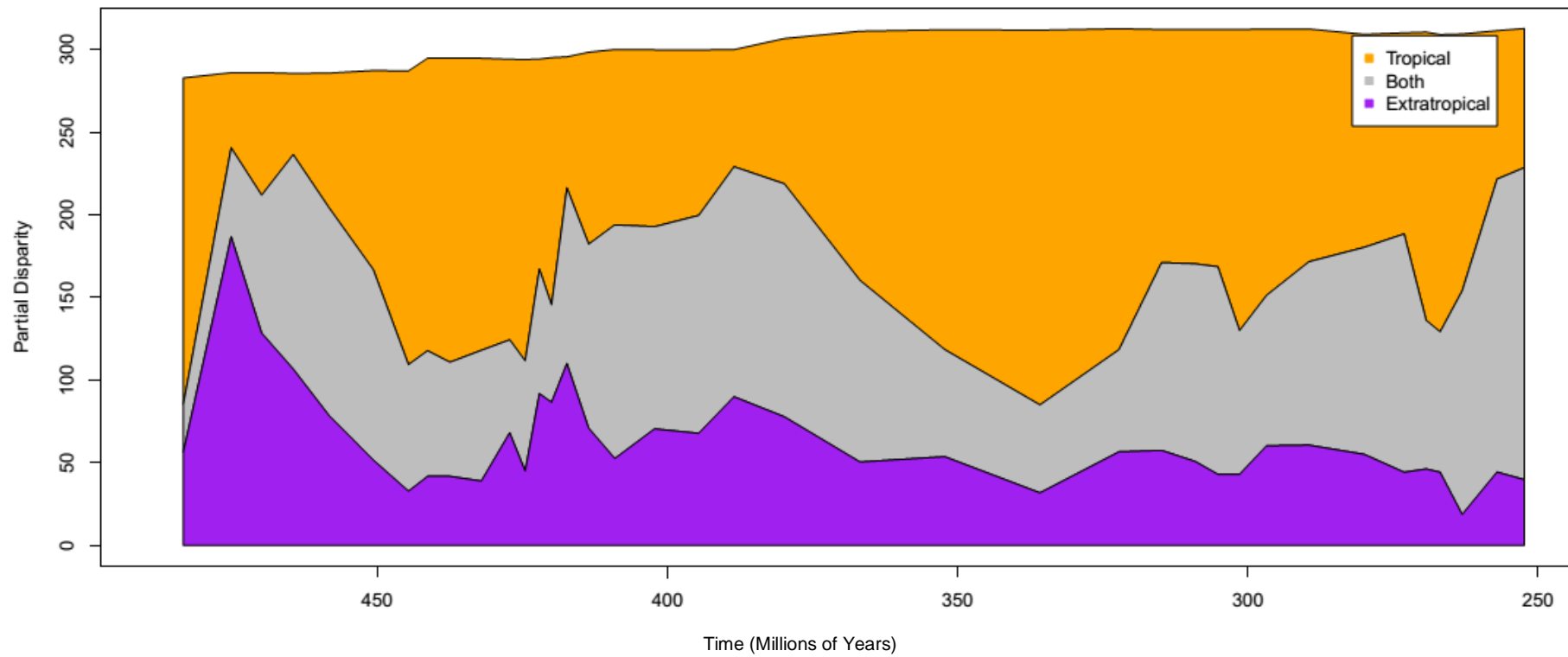


Figure.4.9: Partial disparity based on latitude subgroup across the Palaeozoic. The Mixed Range treatment was used. The height of each colour represents the contribution to the overall disparity in each bin time interval (stages). Orange: Tropical only clades; purple: extra-tropical clades; grey: clades found in both the tropics and extra-tropics.

The total occupation of these groups in morphospace does not change amongst these two subgroups. This suggests that the reduction came from a drop of genera somewhere within the space already occupied by the cosmopolitan genera rather than a rotation or movement of genera at the edges of the space. Through the rest of the carboniferous and Permian, cosmopolitan genera increased in their contribution gradually (Fig.4.9). This gradual increase is punctuated by several drops which occur in tandem with increased contributions from the tropical only group. For most of the increase there is now change in the overall space covered by the three clades, which all share similar amounts of morphospace (Appendix.3). The exception is in the late Permian where there is a drop in the contribution by extra-tropical taxa that is associated with a drastic drop in the space occupied by the extra-tropical genera.

#### **4.10 Discussion**

The results of this Chapter (Fig.4.4-4.9) provide the first attempt to quantify the disparity of bivalves across latitude over the Palaeozoic. They are also the first attempt to quantify latitudinal disparity over such a wide timeframe (the Ordovician – Permian) compared to the study of Grey *et al.* (2010) which only looked at a specific interval (Tithonian). The analyses here also attempted to provide a general account for diversity across latitude for bivalves in the Palaeozoic. The results showed that regardless which gap regime was employed the account of latitudinal diversity across the Palaeozoic was similar to the one produced for the whole marine biota (Fig.4.0) (Vinarskia *et al.* 2011), i.e., that latitudinal diversity gradient maximum rose from the southern latitudes northward towards the equator through the Palaeozoic. Disparity did not follow this pattern, except in the most generalist sense, i.e., disparity values grew northward with diversity through time, which corresponds to the rotation, collision and northward shift of the continents forming Pangaea at the end of the Permian.

Disparity followed a very stable and unchanging pattern across latitude (Fig.4.4AC). One explanation for the unchanging pattern is methodological. The results of Figure 4.7A found that many of the genera that made up the data were defined as cosmopolitan, i.e. genera with latitudinal ranges that ranged through the tropical (0-30°) and extra-tropical (30-90°) latitudinal bins at some point in their evolutionary history. Tropical only genera were more restricted, only those genera that never ventured out of the range of 30°South and 30°North could fall into this category, while in extra-tropical clades the genera could had to be isolated to 30-90°North or South. As the gap regime used allowed for the migration of bivalves through time, if migration across was a prominent feature of bivalves, then they would likely fall into the cosmopolitan category. Genera could also fall into the cosmopolitan category if they straddled the edge of the zone defined here as the tropics, i.e. if they were only found in the middle latitudes (20-40°North or South) then they would still be classed as cosmopolitan.

Future refinements to the method of latitudinal disparity should include splitting the cosmopolitan subgroup either into those occupying the northern or southern hemispheres and/or identify those clades that are isolated at particular latitudes across their entire stratigraphic range. If being cosmopolitan as defined here was a large biasing factor (i.e., disparity is the same across latitude because all genera are found everywhere) it would be expected to have the greatest share of the partial disparity pattern; however, this is not the case (Fig.4.9), and even in some intervals makes up the smallest share of partial disparity (Fig.4.9). This leads to the conclusion that despite the bias caused by the all encompassing nature of cosmopolitan category, the results do not find a large bias towards this subgroup. The partial disparity patterns picked up are interesting as they show that despite total disparity remaining the same across latitude, the taxa making up this increase are a complex mix of tropical only, extra-tropical only and cosmopolitan (or at least migrating but latitudinally isolated) genera.

Recent work on Cenozoic bivalves by Jablonski *et al.* (2013) stresses the importance of 'bridge species' in setting up latitudinal diversity gradients; that is, a species which straddles the line between tropical and extra-tropical latitudes, subsequently carrying their genera out to the tropics. Unfortunately, since my analysis uses genera, identifying Palaeozoic bridge species is beyond the scope of this study, yet it is certainly an avenue for future work. In their paper, Jablonski *et al.* (2013) suggest that bridge species have expanded thermal ranges, originating in the 'core' tropics, rather than evolving at the edges of the tropical range. They do not explicitly mention the range of 'core' tropical latitudes but I assume it to be within the same range as categorised herein (30°South to 30°North).

Bridge species will certainly make up part of the cosmopolitan category. For the majority of the Palaeozoic, tropical only genera contributed the most to the total disparity pattern. When their contribution dropped, the cosmopolitan genera filled the gap. This may suggest that cosmopolitan genera are less disparate than tropical genera, at least in the Palaeozoic. At the end of the Permian cosmopolitan genera do in fact contribute the most to partial disparity which hints at a shift across the boundary and different dynamic in the Mesozoic and Cenozoic. Nevertheless, until analyses like those herein are carried toward these other time intervals this can only be speculation.

Studies have suggested that the tropics are a macroevolutionary source, a place where diversity can be generated which supplies the extra-tropical latitudes with diversity (Jablonski *et al.* 2006; Krug *et al.* 2009). The extra-tropical latitudes have also been identified as less extinction prone, at least of specific catastrophic events (Vilhena *et al.* 2013). For disparity there does not seem to be a gradient in the Palaeozoic. The exception is the Ordovician-Silurian. This gradient is unlikely to correspond to a gradient of disparity decrease caused by genera having similar morphologies. Instead this gradient it may be explained by bivalves evolving in the southern hemisphere (Sanchez 2008) and

not yet reaching the high northern latitudes (most of the northern latitudes sediments in the Early Palaeozoic do not have bivalve fossils).

Given bivalves are found across the globe today, and that the majority of Palaeozoic ocean rock has been subducted since the Permian, the pattern of low diversity and disparity in the northern latitudes could be the result a major gap in the fossil record. There is a question of the level of accuracy of the results given this caveat. The diversity of other benthic taxa, such as Terebratulida brachiopods (Fitzgerald & Carlson 2006), echinoderms (Markov *et al.* 2012), and the marine biota as a whole (Vinarskia *et al.* 2011) do not show diversity estimates for these northern palaeolatitudes. As the results of this Chapter are shared in other clades, one can interpret them in two ways. Either the low diversity in the northern palaeolatitudes is a real pattern or there is a fossil record bias spanning the much of the Palaeozoic that has affected all marine clades alike. Unfortunately there may be no true way to be sure which of these explanations is more correct. The attempt at finding the answer would be to find as many Palaeozoic marine fossil sites as possible belonging to northern palaeolatitudes and to make a concerted effort to assess the quality of those fossil sediments (e.g. level of taphonomic bias).

If the tropics are indeed a macroevolutionary source (Jablonski *et al.* 2006) which has persisted since the Cambrian (Vinarskia *et al.* 2011), the stability of bivalves across the Palaeozoic may be due to a gradual increase in morphological diversity migrating from the tropical latitudes into the higher latitudes. Nevertheless, the degree of disparity across latitudes could also be levelled out by some genera keeping a foothold in the tropics (Jablonski *et al.* 2006; Krug *et al.* 2009). The best way to test whether disparity was indeed a result of genera in the Palaeozoic extending their ranges towards the tropics would involve looking at a phylogenetic tree that could be trusted and measuring disparity along its branches (i.e. is disparity distributed across families or within them) and then linked to the latitudinal distribution of these genera (Harmon *et al.* 2003).

By tracking the changes in disparity and latitude along a phylogenetic tree of Palaeozoic bivalve genera one begin to tease apart the contribution of tropical and extra-tropical genera to the latitudinal disparity pattern that accounts for similarities in disparity due to shared ancestry alone. As cosmopolitan genera do not make up the majority of disparity contribution this suggests that for much of the Palaeozoic, the major contribution is by tropical only genera, giving more weight to the idea of the tropics as a macroevolutionary sources (in this case for disparity).

#### **4.11 Conclusion & Future Prospects**

The main aim of this chapter was to quantify the disparity of Palaeozoic bivalves across latitude to address how disparity itself changes with latitude. The results show that latitudinal disparity did not follow a gradient for the much of the Palaeozoic, the exception being the Early Palaeozoic (Ordovician-Silurian) where diversity in northern latitudes was

low. This is contrary to the latitudinal diversity data that followed a gradient similar to that found for the marine biota and echinoderms, brachiopods, molluscs and the marine biota as a whole (Fitzgerald & Carlson 2006; Vinarskia *et al.* 2011), i.e., a peak in the tropics, tapering off toward the poles. This shows that there was decoupling between diversity and disparity across latitude for most of the Palaeozoic. The reasons behind the stable disparity pattern could be methodological. The data used here was (and still is) the most detailed morphological dataset available that includes a wide selection of Palaeozoic bivalves, but if the time intervals selected are too coarse and the disparity pattern subtle, then any patterns across time may have been lost. Finer scale analyses in individual periods will hopefully pick out any changes in disparity that are at a finer scale than the one used here (genera, and stages).

Another possibility is that the indifference of disparity to change across latitude was a real pattern. The disparity vs. latitude data of Grey *et al.* (2010) suggests a flat trajectory across latitude for *Buchia* in the Tithonian. Scaling that singular result to encompass all bivalve genera across the Phanerozoic would be unwise. Despite the stability of disparity across palaeolatitudes, the contribution of tropical vs. extra-tropical vs. cosmopolitan genera is dynamic and an area of potential future research. One project could explore the cosmopolitan category, identifying the disparity of supposed bridge species (Jablonski *et al.* 2013) against those that are tropical only to see if there were differences in disparity specifically between the two groups, i.e. are tropical only genera more or less disparate than those that bridge the gap between the tropics and extra-tropics? The tropical only genera herein had a relatively stable disparity trajectory through time. Tropical only taxa have larger latitudinal ranges due to them occupying narrow but far reaching temperature ranges (Jablonski *et al.* 2013) while higher latitude taxa tend to have shorter range limits, even though they may have a more tolerant thermal range. This begs the question whether the disparity of bivalves is at least partly the consequence of this pattern in thermal tolerance.

An obvious next step is to pick individual periods of time - the mass extinction events of the End Ordovician and Late Devonian being logical choices due to their putative change in diversity across latitude particularly in the latter extinction (Fig.4.4). Putative peaks in the Silurian (Fig.4.4, and Chapter 3), combined with the fluctuating contribution of the three latitudinal subgroups, suggests this would be another time period worth closer inspection. Adding Mesozoic and Cenozoic genera to the data used here would facilitate a latitudinal disparity curve across the Palaeozoic. This could subsequently verify the increased contribution of cosmopolitan genera in the latest Permian as a paradigm shift in the role of tropical only/cosmopolitan genera in bivalve disparity and diversity.

The movement of the continents seem to have had a profound effect on the location of the peak in the bivalve latitudinal diversity gradient, but not so much once

northern and southern latitudes were open to bivalve diversification and migration (Fig.4.4). There may still be differences in disparity when comparing geographical regions, i.e. does disparity track provinciality in the Palaeozoic? The southern hemisphere Gondwanan basins of Argentina have been suggested to be a source of diversity (Sanchez 2008), but the results here suggest disparity is stable across latitude. There is still the question of whether disparity is sourced from different areas (i.e. hotspots of morphological diversification), and so looking at different provinces through time may help to distinguish any hotspots if they exist. There may also be the issue of collectorship. Any identified hotspots such as the one claimed by Sanchez (2008) may be an artefact of increased sampling and focus on those specific areas. To identify any “true” hotspots of biodiversity future work should focus on assessing diversity and disparity trends in a number of regions across the globe simultaneously.

## ***Conclusions & Future Work***

The primary objectives of the thesis were to quantify and systematically test how disparity has evolved throughout the Phanerozoic across the Metazoa, and then to focus on an individual clade to assess how disparity changed through time and across latitude. The following sections provide the main conclusions of this thesis and the possible future directions others may wish to explore using the data or techniques described in this thesis.

### ***Early High Disparity: the Rule***

The analyses of Chapter 1 found that early high disparity does indeed appear to be the rule as it was the most prevalent disparity pattern amongst the Metazoa across the Phanerozoic; however, the results did not show any monotonic trend in disparity through the Phanerozoic. Mass extinctions events were the exception as the majority of clades terminating at them had a top-heavy disparity profile. The lack of temporal trend in disparity through the Phanerozoic suggests that the processes governing the formation of one disparity profile over another have not changed over geological time. If early high disparity patterns were mainly the result of ecospace filling (Valentine 1995) then one may expect more bottom-heavy clade profiles in the wake of mass extinctions. Surprisingly the results show no preference for that kind of disparity pattern. The results also show that there was no transition in disparity profile patterns leading up to or between mass extinctions.

As with all analyses, there are some improvements and future directions that could be explored to better understand the changes in disparity across the Phanerozoic and at mass extinctions. As the aim of the Chapter was to assess disparity across the Phanerozoic, the morphological datasets of clades originating directly after mass extinction events were few. The results finding no preference for bottom-heavy profiles in the wake of mass extinctions should then be treated as preliminary as the addition of more clades originating in the wake of mass extinctions could change the results significantly. To counter this one could begin to analysing clades that begin to rapidly rise in diversity shortly after mass extinctions such as the mammals after the Cretaceous impact event or bivalves after the Permian mass extinction. Chapter 1 also tried to identify transitions from one disparity pattern to another between mass extinctions but needed to pool data as the numbers of clades between individual extinction events were too few. By focusing on clades between specific extinction events one could begin to more rigorously test whether disparity patterns change from one mass extinction event to another. In addition to focusing on clades ending at specific events, another step would be to look at different kinds of morphological data through the Phanerozoic. Morphometric data like outlines and landmarks explore different aspects of shape. Comparing the results found in



this thesis to results based on morphometric data would provide more complete picture of how disparity has evolved across the Phanerozoic.

There is the possibility that early high disparity in most clades is the result of an exhaustion of possible discrete characters (Wagner 2000; Wagner *et al.* 2006). The exhaustion of characters does not in itself suggest a methodological bias unless on average those authors who build cladistic or phenetic datasets code their data in the same general way (unlikely given the radical differences in morphology between some groups). Future work should focus on exploring the role that character exhaustion plays in shaping disparity patterns in combination using disparity patterns based on morphometric data. In that way one could begin to identify how many clades showing early high disparity based on discrete characters is true biological variation and not an artefact of authors selecting a small subsample of possible morphological characters. The script used in the analyses herein focused on mainly on the sum of variances as the measure of disparity. Adding in other disparity measures will provide additional information about the exploration of clades in morphospace (Ciampaglio *et al.* 2001).

Another logical next step from the work of this thesis would be to look at the individual clades. By looking at individual clades, an alternative form of disparity analysis could be considered using the data of Chapter 1: do clades that share certain disparity profiles partition this disparity along their phylogenetic tree in similar ways (Harmon *et al.* 2003)? The data of Chapter 1 did not attempt to include ghost range data. Producing phylogenetic trees of each dataset, and combining this information with the stratigraphic record could help to more accurately identify the timings of highest disparity, and identify whether disparity in each clade is partitioned differently depending on the kind of disparity profile. Subgroups contribute to disparity in very different ways through time as seen in Chapters 3 and 4 and in the literature (Foote 1993a) but discrete character partial disparity has never been tracked along the phylogenetic trees. Another aspect of the effect of mass extinctions on Phanerozoic disparity that could not be address in Chapter 1 due to its broad scale nature is the effect that mass extinctions had on those clades that survived them. The work of Chapter 3 found that mass extinction in bivalves removed genera that on the whole were less disparate than the survivors, while the work of Chapter 1 showed that clades terminating at mass extinctions mainly followed top-heavy disparity profiles. It is certainly unclear whether the increased disparity effect of mass extinctions on clades as shown in Chapter 3 are the general rule across all clades and across all mass extinction events. The data of Chapter 1 provides a healthy sample with which to begin to explore the way in which mass extinctions affect the disparity patterns of survivor clades.

## ***Bivalve Diversity Across the Phanerozoic***

The results of Chapter 2 found that the bivalve diversity pattern through the Phanerozoic seems to be relatively robust to the method of diversity calculation. This is a promising result given that the level of stratigraphic revisions and new techniques available to calculate diversity and remove biases e.g. shareholder quorum subsampling and sampling proxy methods. This would suggest that despite potential biases, current understanding of the bivalve fossil record is robust. An interesting consequence of this is that bivalve diversity has been growing since the Cambrian, and has yet to flatten out. The next step should be identifying the reasons why bivalves have been as successful through the Phanerozoic as a whole, and why despite the number of mass extinctions, bivalves have not shown any signs of slowing in diversity. An aspect of diversity studies not conducted here but should be an area of future research is the revision of bivalve origination and extinction rates across the Phanerozoic, using the updated ranges of the data herein and the revised Treatise of Invertebrate Palaeontology.

One clade is missing from the Treatise data herein: the Hippuritoida (rudist bivalves). These data should be included in future revisions of the analyses conducted herein. These bivalves could not be included during the analyses and writing of this thesis as the clade was still undergoing revisions. Their impact on the diversity pattern of the Phanerozoic will not be profound to as they were a Mesozoic only clade; however, their inclusion would likely have some effect on the story of bivalves in the Mesozoic. Another possible area of future work would be to explore the effects of other fossil record proxies on the diversity curve of bivalves such as outcrop area, or country of origin or removal of lagerstattë. These other sampling proxies could relate issues not picked up in the data here, such as issues relating to the areas in which the fossils are found. Some outcrop areas will have provided more diversity because those outcrop areas were larger or better explored and so smaller outcrop areas may have much diversity still hidden. If many of the fossils found so far are biased to particular areas of the globe (e.g. North America, Europe) then the signal being picked up could potentially not be a global signal of diversity increase but a regional one, and requiring its own set of explanations i.e. why could diversity in a region continue to increase across the Phanerozoic? The best way to counter this possible bias would be to identify the areas of the globe not represented by the current set of fossils and include more specimens from areas those areas. Finally lagerstattë could be a problem for the inference of important diversity peaks through the Phanerozoic, and so the data from these types of deposit should be removed to see what peaks in diversity are the results of this issue.

An essential task will be to merge the data in the Treatise with that of the Sepkoski and Paleobiology Databases. Although having multiple resources of stratigraphic data on the surface is good for debating the correct stratigraphic ranges of fossils, in the case of the bivalve ranges found in the Sepkoski, revised Treatise, and Paleobiology Database,

all originate from the same original source (Cox *et al.* 1971). The differences do not seem to affect the overall trajectory of bivalves, but for getting at the reasons behind particular peaks and troughs, a single, continually updated and agreed upon source of diversity data would be more useful than the several data sources differing in the timing of particular events. Another logical step will be to look in detail at the reasons behind individual events both in the bivalves as a group but also for the various subgroups (Protobranchia, Heteroconchia and Pteriomorphia).

### ***Disparity in Bivalves Increases Across the Palaeozoic***

Chapter 3 wished to address the disparity changes in bivalves through the Palaeozoic, and found that disparity increases with time across the Palaeozoic when bivalves are considered as one single clade. The contribution of different subclades to the overall disparity trajectory has changed through that time. Protobranchia made its biggest contribution in the Ordovician, and slowly dropped in importance (in disparity and diversity terms) through the rest of the Palaeozoic. The contribution by Heteroconchia remained stable for much of the Palaeozoic, growing in the Devonian but then remaining stable for the rest of the Palaeozoic. Pteriomorphia continued to grow in disparity throughout the Palaeozoic and explored the largest portion of the morphospace. By the end of the Permian Pteriomorphia made up the largest contribution of disparity. Each subclade varied in their pattern of morphospace clustering through the Palaeozoic. For Protobranchia there was a gradual drop, in Pteriomorphia clustering increased while in Heteroconchia clustering fluctuated but did not trend in any direction. The clustering results suggest that as the clades evolved, Pteriomorphia changed from exploring the whole space to clustering within particular regions, Heteroconchia explored within their own space though the whole Palaeozoic while Protobranchia remained clustered within two regions but in low abundance by the end of the Permian.

There was disconnection between the diversity and disparity of bivalves over the Palaeozoic. Diversity in bivalves increased in the Ordovician-Silurian, dropping thorough the Devonian and Carboniferous and growing in the Permian before the End Permian mass extinction. During that time the disparity of Bivalvia the three subclades when studies in isolation increased with no significant jumps or drops in disparity. The modelling approach found that each clade followed an Ornstein-Uhlenbeck process of disparity increase across the Palaeozoic. This process has been used in the past for suggesting adaptive radiations based upon optimisation in particular characters; however, for disparity the process could be explained by taxa reaching some boundary stopping them from occupying new environmental niches. The modelling approach of Hunt (2006) was initially designed to model the evolution of singular characters/characters states. The analyses of Chapter 3 were suitable as a first attempt to use the techniques to understand the processes behind disparity changes in clades; however, much more work is needed in

this area. As a first step the approach of taking the existing method and applying it to disparity was logical; however, the next step will be to assess whether the normal distributions used in the parameterisation of these models is correct for disparity data. Further work will be needed to adjust the meaning of the current models (adaptive radiation, stasis, genetic drift) to take into account that the disparity of clades is being modelled rather than individual characters.

As for the non-linear properties observed in the morphological data of Chapter 3 (Cailliez corrected vs. uncorrected), future work may wish to explore the reasons behind them and whether they affect the chosen model of disparity. It would also be prudent to look for similar non-linear properties using techniques used in this thesis in other published disparity analyses as a way of confirming whether the axes selection issues highlighted Chapter 3 are bivalve specific or a much wider issue. Another area of potential future research is the analysis of bivalve disparity using other methods. Once a tree of the genera used in this Chapter becomes available disparity could be calculated along the tree to distinguish whether the morphological variability is distributed within particular families or is evenly distributed across the all subclade families. Combined with information on their stratigraphic range and latitudinal distribution one could attempt to pinpoint the sources (e.g. geographical provinces, environmental conditions, habit use, taxonomic grouping) encouraging bivalve morphological variability.

The marginal likelihood analyses of Chapter 3 also provided an interesting result when mass extinction survivors vs. terminated clades were compared. In the End Ordovician and Late Devonian mass extinctions there was an increase in disparity across the boundaries. This suggests that as well as being times of mass death these events also helped to increase the disparity of bivalves that survived. The analyses concerning this result are only preliminary, requiring further refinement. For example the analyses did not take into account the relationship between those who survived vs. those that did not. To better assess if there was an increase in disparity across the boundary will require assessing the disparity in related taxa across the boundaries. Using the techniques found in this thesis in other clades will also be important to address whether all mass extinctions increase disparity in survivors or whether clades disparity profiles are affected in different ways. Finding the answer to either of these will provide greater insight into how biodiversity is affected by mass extinctions.

### ***Latitudinal Disparity in Bivalves Remains Stable Across the Palaeozoic***

The result of Chapter 4 found that disparity in bivalves remained stable across latitude for the majority of the Palaeozoic. The exception to this is the gradient that is apparent in the Early Palaeozoic (Ordovician-Silurian). It is likely that this result is caused by the reduced number of taxa in those latitudes rather than lower levels of disparity. Future work could confirm this by looking at bivalve disparity in this period in more detail.

One way to go about this is to assess the disparity of as many bivalve rich fossils localities of different latitudes as possible across this time frame. Diversity of bivalves across latitude was very different to latitudinal disparity. Like the results of other Palaeozoic marine clades bivalve latitudinal diversity peaked in the southern hemisphere and travelled north through the Palaeozoic as the continents rotated and collided to form the supercontinent Pangaea by the end of the Permian. The disparity contributed by tropical, extra-tropical and cosmopolitan groups was dynamic across the Palaeozoic and presents an area for future research. The tropics are suggested to feed the poles with diversity while retaining the diversity they generate, making them both a museum and source of diversity. By adding phylogenetic data of the genera of this Chapter one could take the latitudinal disparity and track how it evolved along the tree, i.e., whether the majority of the disparity is contributed by particular lineages, or distributed evenly. Bridge species have been proposed as important in setting up and continuing the diversity gradient across latitude over time (Jablonski *et al.* 2013). One potential project would be to assess their importance in terms of disparity, i.e. are bridge species more disparate than tropical only or extra-tropical only clades? If true then it would provide evidence that the tropics or the edges of the tropics are a source of increased disparity.

A future direction of the work here would be to begin to include the morphological data across a range of Mesozoic and Cenozoic bivalves to build up the picture of latitudinal (and 'regular') disparity through the Phanerozoic. The work here has taken a large first step in making bivalves a model clade for disparity and diversity across time and latitude which currently does not exist. From the construction of this dataset future researchers could if they so wish be able to use bivalves to identify the changes in morphological disparity across the globe for the other mass extinctions in the Phanerozoic. This is important given the idea that morphological variation and higher latitudinal placement may buffer organisms to extinction (Kolbe *et al.* 2011; Vilhena *et al.* 2013).

Understanding why disparity does not seem to fluctuate with latitude should be looked at in more detail, given that latitudinal diversity trends are a prominent feature of the fossil record. The analyses here identified that the Devonian was a time where Heteroconchia were rising in partial disparity, while latitudinal diversity (and diversity of Bivalvia in general) was decreasing. There is the possibility that different subclades were isolated to particular latitudinal provinces caused by particular tolerances to ocean temperature (Jablonski *et al.* 2013). Individual projects covering particular periods would be the best compromise between the broad scale approach of Chapter 4 and the fine scale approach defined in Grey *et al.* (2010).

Overall, the research in this thesis has quantified the disparity of metazoans across the Phanerozoic, updated the account of bivalve diversity through the Phanerozoic and taken

the first step of establishing bivalves as a prominent clade for the analysis of disparity across time and latitude. It is my hope that those who have read this thesis found this research and techniques useful in focusing their own ideas about the evolution of biodiversity across time and space, and that maybe they will look at bivalves as more than just some shells found on the beach, or dinner.

## References

- Aberhan, M., Nurnberg, S. & Kiessling, W., 2012. Vision and the diversification of Phanerozoic marine invertebrates. *Paleobiology*, 38(2), pp.187–204.
- Adamowicz, S.J., Purvis, A. & Wills, M.A., 2008. Increasing morphological complexity in multiple parallel lineages of the Crustacea. *Proceedings of the National Academy of Sciences of the United States of America*, 105(12), pp.4786–4791.
- Adnet, S. & Cappetta, H., 2001. A palaeontological and phylogenetical analysis of squaliform sharks (Chondrichthyes: Squaliformes) based on dental characters. *Lethaia*, 34(3), pp.234–248.
- Adrain, J.M., Chatterton, B.D.E. & Kloc, G.J., 2008. Systematics of the koneprusiine trilobites, with new taxa from the Silurian and Devonian of Laurentia. *Journal of Paleontology*, 82(4), pp.657–675.
- Allen, A.P. & Gillooly, J.F., 2006. Assessing latitudinal gradients in speciation rates and biodiversity at the global scale. *Ecology Letters*, 9(8), pp.947–954.
- Alroy, J. et al., 2001. Effects of sampling standardization on estimates of Phanerozoic marine diversification. *Proceedings of the National Academy of Sciences of the United States of America*, 98(11), pp.6261–6266.
- Alroy, J., 2010a. Fair sampling of taxonomic richness and unbiased estimation of origination and extinction rates. In *Quantitative Methods in Paleobiology*. pp. 55–80.
- Alroy, J., 2010b. Geographical, environmental and intrinsic biotic controls on Phanerozoic marine diversification. *Palaeontology*, 53(6), pp.1211–1235.
- Alroy, J. et al., 2008. Phanerozoic trends in the global diversity of marine invertebrates. *Science*, 321(5885), pp.97–100.
- Alroy, J., 2010c. The shifting balance of diversity among major marine animal groups. *Science*, 329, pp.1191–1194.
- Alvarez, F. et al., 1998. The classification of athyridid brachiopods. *Journal of Paleontology*, 72(5), pp.827–855.
- Anderson, D.R., Burnham, K.P. & Thompson, W.L., 2000. Null hypothesis testing: problems, prevalence, and an alternative. *Journal of Wildlife Management*, 64(4), pp.912–923.
- Anderson, L.C. & Roopnarine, P.D., 2003. Evolution and phylogenetic relationships of Neogene Corbulidae (Bivalvia; Myoidea) of tropical America. *Journal of Paleontology*, 77(6), pp.1086–1102.
- Anderson, L.C., Wesselingh, F.P. & Hartman, J.H., 2010. A phylogenetic and morphologic context for the radiation of an endemic fauna in a long-lived lake: Corbulidae (Bivalvia; Myoidea) in the Miocene Pebas Formation of western Amazonia. *Paleobiology*, 36(4), pp.534–554.
- Anderson, P.S.L. et al., 2011. Initial radiation of jaws demonstrated stability despite faunal and environmental change. *Nature*, 476(7359), pp.206–209.

- Anstey, R.L. & Pachut, J.F., 1995. *Phylogeny, diversity history, and speciation in Paleozoic bryozoans* D. H. Erwin & R. L. Anstey, eds., New York: Columbia University Press.
- Archie, J.W., 1989a. A Randomization test for phylogenetic Information in systematic data. *Systematic Zoology*, 38(3), pp.239–252.
- Archie, J.W., 1989b. Homoplasy excess ratios - new indexes for measuring levels of homoplasy in phylogenetic systematics and a critique of the consistency index. *Systematic Zoology*, 38(3), pp.253–269.
- Archie, J.W., 1996. *Homoplasy: the recurrence of similarity in evolution* M. J. Sanderson & L. Hufford, eds., Waltham, MA: Academic.
- Arita, H.T. & Vazquez-Dominguez, E., 2008. The tropics: cradle, museum or casino? A dynamic null model for latitudinal gradients of species diversity. *Ecology Letters*, 11(7), pp.653–663.
- Arthur, W., 2006. *Creatures of accident: the rise of the animal kingdom*, New York: Hill & Wang.
- Arthur, W., 2004. The effect of development on the direction of evolution: toward a twenty-first century consensus. *Evolution & Development*, 6(4), pp.282–288.
- Ausich, W.I. & Deline, B., 2012. Macroevolutionary transition in crinoids following the Late Ordovician extinction event (Ordovician to Early Silurian). *Palaeogeography, Palaeoclimatology, Palaeoecology*, 361, pp.38–48.
- Bambach, R.K., Knoll, A.H. & Wang, S.C., 2004. Origination, extinction, and mass depletions of marine diversity. *Paleobiology*, 30(4), pp.522–542.
- Bapst, D.W. et al., 2012. Graptoloid diversity and disparity became decoupled during the Ordovician mass extinction. *Proceedings of the National Academy of Sciences of the United States of America*, 109(9), pp.3428–3433.
- Barrett, P.M. & McGowan, A.J., 2009. Dinosaur diversity and the rock record. *Proceedings of the Royal Society B*, 276, pp.2667–2674.
- Benton, M.J. et al., 2013. The first half of tetrapod evolution, sampling proxies, and fossil record quality. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 372, pp.18–41.
- Benton, M.J., 1993. The fossil record 2. *Chapman & Hall, London*, p.845.
- Benton, M.J., 2009. The red queen and the court jester: species diversity and the role of biotic and abiotic factors through time. *Science*, 323(5915), pp.728–732.
- Benton, M.J., 2000. *Vertebrate palaeontology* Second., Oxford: Blackwell Science.
- Benton, M.J. & Emerson, B.C., 2007. How did life become so diverse? The dynamics of diversification according to the fossil record and molecular phylogenetics. *Palaeontology*, 50(1), pp.23–40.
- Bertelli, S., Giannini, N.P. & Ksepka, D.T., 2006. Redescription and phylogenetic position of the early Miocene penguin *Paraptenodytes antarcticus* from patagonia. *American Museum Novitates*, 3525(1), pp.0–36.



- Bervoets, F., 2011. DinoData. Available at: [www.dinodata.org/](http://www.dinodata.org/) [Accessed January 17, 2011].
- Blakey, R., 2011. Library of Paleogeography. Available at: <http://cpgeosystems.com/paleomaps.html> [Accessed February 9, 2012].
- Body, R.L., 2000. Index of Mollusks. *The Nature Mapping Program*. Available at: <http://naturemapping> [Accessed September 16, 2011].
- Bouchet, P. et al., 2010. Nomenclator of bivalve families with a classification of bivalve families. *Malacologia*, 52(2), pp.1–184.
- Brake, I., 2011. Fungus Gnats Online. Available at: <http://sciarioidea.info/> [Accessed April 8, 2011].
- Brands, S.J., 2011. Systema Naturae 2000: The Taxonomicon. Available at: <http://taxonomicon.taxonomy.nl/> [Accessed February 21, 2011].
- Briggs, D.E.G., Fortey, R.A. & Wills, M.A., 1992. Morphological disparity in the Cambrian. *Science*, 256(5064), pp.1670–1673.
- Brusatte, S.L. et al., 2012. Dinosaur morphological diversity and the end-Cretaceous extinction. *Nature Communications*, 3(803), pp.1–8.
- Brusatte, S.L. et al., 2008. The first 50 Myr of dinosaur evolution: macroevolutionary pattern and morphological disparity. *Biology Letters*, 4(6), pp.733–736.
- Burbrink, F.T. & Pyron, R.A., 2010. How does ecological opportunity influence rates of speciation, extinction, and morphological diversification in New World ratsnakes (tribe Lampropeltini)? *Evolution*, 64(4), pp.934–943.
- Butler, M.A. & King, A.A., 2004. Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *The American Naturalist*, 164(6), pp.683–695.
- Butler, R.J. et al., 2009. Estimating the effects of sampling biases on pterosaur diversity patterns: implications for hypotheses of bird/pterosaur competitive replacement. *Paleobiology*, 35(3), pp.432–446.
- Cailliez, F., 1983. The analytical solution of the additive constant problem. *Psychometrika*, 48(2), pp.305–308.
- Carlson, S.J., 1992. Evolutionary trends in the articulate brachiopod hinge mechanism. *Paleobiology*, 18(3).
- Carlson, S.J. & Fitzgerald, P.C., 2007. Sampling taxa, estimating phylogeny and inferring macroevolution: an example from Devonian terebratulide brachiopods. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh*, 98(Special Issue 3-4), pp.311–325.
- Carroll, R.L., 1988. *Vertebrate paleontology and evolution*, New York: W. H. Freeman and Company.
- Carter, J.G. et al., 2011. A synoptical classification of the Bivalvia (Mollusca). *Paleontological Contributions*, 4, pp.1–49.

- Carter, J.G., Campbell, D.C. & Campbell, M., 2000. Cladistic perspectives on early bivalve evolution. In *The Evolutionary Biology of the Bivalvia*. Geological Society, London, Special Publications, pp. 47–79.
- Carter, J.G., Campbell, D.C. & Campbell, M., 2006. Morphological phylogenetics of the early Bivalvia. In M. Malchus & J. M. Pons, eds. *International Congress on Bivalvia*.
- Caruthers, A.H., Smith, P.L. & Gröcke, D.R., 2013. The Pliensbachian–Toarcian (Early Jurassic) extinction, a global multi-phased event. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 386, pp.104–118.
- Chipman, A.D., 2001. Developmental exaptation and evolutionary change. *Evolution & Development*, 3(5), pp.299–301.
- Chun-lian, L., 2000. Extinction events among Jurassic bivalves. *Acta Scientiarum Naturalium Universitatis Sunyatseni*, 39(1), pp.91–95.
- Ciampaglio, C.N., Kemp, M. & McShea, D.W., 2001. Detecting changes in morphospace occupation patterns in the fossil record: characterization and analysis of measures of disparity. *Paleobiology*, 27(4), pp.695–715.
- Cisneros, J.C. & Ruta, M., 2010. Morphological diversity and biogeography of procolophonids (Amniota: Parareptilia). *Journal of Systematic Palaeontology*, 8(4), pp.607–625.
- Clack, J.A., 2006. The emergence of early tetrapods. *Palaeogeography Palaeoclimatology Palaeoecology*, 232(2-4), pp.167–189.
- Clauset, A. & Erwin, D.H., 2008. The evolution and distribution of species body size. *Science*, 321(5887), pp.399–401.
- Colwell, R.K. & Coddington, J.A., 1994. Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions Of The Royal Society Of London. Series B Biological Sciences*, 345(1311), pp.101–118.
- Cope, J.C.W., 2000. A new look at early bivalve phylogeny. In E. M. Harper, J. D. Taylor, & J. A. Crame, eds. *The Evolutionary Biology of the Bivalvia*. London: Geological Society, London, Special Publications, pp. 81–95.
- Cope, J.C.W., 2004. Bivalve and rostroconch mollusks. In B. D. Webby et al., eds. *The Great Ordovician Biodiversification Event*. New York: Columbia University Press, pp. 196–208.
- Cope, J.C.W., 2002. Diversification and biogeography of bivalves during the Ordovician Period. *Geological Society, London, Special Publications*, 194, pp.35–52.
- Cope, J.C.W., 1999. Diversification of bivalves in the Ordovician. *Geobios*, 32(2), pp.175–185.
- Cope, J.C.W., 1996. Early Ordovician (Arenig) Bivalves from the Llangynog inlier, south Wales. *Palaeontology*, 39(4), pp.979–1025.
- Cox, L.R. et al., 1971. Bivalvia. In R. C. Moore, ed. *Treatise of Invertebrate Paleontology, Part N, Mollusca* 6. The Geological Society of America and University of Kansas, p. 1224.

- Crame, J.A., 2002. Evolution of taxonomic diversity gradients in the marine realm: a comparison of Late Jurassic and Recent bivalve faunas. *Paleobiology*, 28(2), pp.184–207.
- Crampton, J.S. et al., 2003. Estimating the rock volume bias in palaeodiversity studies. *Science*, 301, pp.358–360.
- Crepet, W.L. & Niklas, K.J., 2009. Darwin's second "abominable mystery": Why are there so many angiosperm species? *American Journal of Botany*, 96(1), pp.366–381.
- Darwin, C., 1859. *On the origin of species by natural selection*, London: Murray.
- Davies, R.G. et al., 2007. Topography, energy and the global distribution of bird species richness. *Proceedings of the Royal Society B*, 274(1614), pp.1189–1197.
- Deline, B. & Ausich, W.I., 2011. Testing the plateau: a reexamination of disparity and morphologic constraints in early Paleozoic crinoids. *Paleobiology*, 37(2), pp.214–236.
- Dera, G. et al., 2010. High-resolution dynamics of Early Jurassic marine extinctions: the case of Pliensbachian–Toarcian ammonites (Cephalopoda). *Journal of the Geological Society*, 167(1), pp.21–33.
- Dmitriev, V.Y. & Nevesskaja, L.A., 2006. Taxonomic diversification of normal-marine bivalve mollusks in the Phanerozoic. *Paleontological Journal*, 40(3), pp.253–260.
- Dommergues, J.-L., Laurin, B. & Meister, C., 1996. Evolution of ammonoid morphospace during the Early Jurassic radiation. *Paleobiology*, 22(2), pp.219–240.
- Donoghue, M.J., 2005. Key innovations, convergence, and success: macroevolutionary lessons from plant phylogeny. *Paleobiology*, 31(2), pp.77–93.
- Donovan, D.A., Elias, J.P. & Baldwin, J., 2004. Swimming behavior and morphometry of the file shell *Limaria fragilis*. *Marine and Freshwater Behaviour and Physiology*, 37(1), pp.7–16.
- Dornburg, A. et al., 2011. The influence of an innovative locomotor strategy on the phenotypic diversification of triggerfish (family: Balistidae). *Evolution*, 65(7), pp.1912–1926.
- Dowton, M. & Austin, A.D., 2002. Increased Congruence Does Not Necessarily Indicate Increased Phylogenetic Accuracy-The Behavior of the Incongruence Length Difference Test in Mixed-Model Analyses. *Systematic Biology*, 51(1), pp.19–31.
- Dumont, E.R. et al., 2012. Morphological innovation, diversification and invasion of a new adaptive zone. *Proceedings of the Royal Society B-Biological Sciences*, 279(1734), pp.1797–1805.
- Duperron, S. et al., 2013. An overview of chemosynthetic symbioses in bivalves from the North Atlantic and Mediterranean Sea. *Biogeosciences*, 10(5), pp.3241–3267.
- Eades, D.C. et al., 2011. Orthoptera Species File Online Version 2.0/4.1. Available at: <http://orthoptera.speciesfile.org> [Accessed May 13, 2011].
- Eble, G.J., 2000. Contrasting evolutionary flexibility in sister groups: disparity and diversity in Mesozoic atelostomate echinoids. *Paleobiology*, 26(1), pp.56–79.

- Edwards, A.W.F., 1992. *Likelihood: the expanded edition*, Baltimore: MD: The Johns Hopkins University Press.
- Efron, B., 1982. *The Jackknife, the bootstrap, and other resampling plans*, Society for Industrial & Applied.
- Elicki, O. & Gursu, S., 2009. First record of *Pojetaia runnegari* Jell, 1980 and Fordilla Barrande, 1881 from the Middle East (Taurus Mountains, Turkey) and critical review of Cambrian bivalves. *Paläontologische Zeitschrift*, 83(2), pp.267–291.
- Erwin, D.H., 2007. Disparity: morphological pattern and developmental context. *Palaeontology*, 50(1), pp.57–73.
- Erwin, D.H., 1994. Early introduction of major morphological innovations. *Acta Palaeontologica Polonica*, 38(3/4), pp.281–294.
- Erwin, D.H., 2011. Evolutionary uniformitarianism. *Developmental Biology*, 357(1), pp.27–34.
- Erwin, D.H., 2001. Lessons from the past: biotic recoveries from mass extinctions. *Proceedings of the National Academy of Sciences of the United States of America*, 98(10), pp.5399–5403.
- Etienne, R.S. & Haegeman, B., 2012. A conceptual and statistical framework for adaptive radiations with a key role for diversity dependence. *American Naturalist*, 180(4), pp.E75–E89.
- Etter, R.J. et al., 2005. Population differentiation decreases with depth in deep-sea bivalves. *Evolution*, 59(7), pp.1479–1491.
- Evenhuis, N.L., 1994. *Catalogue of the fossil flies of the world (Insecta: Diptera)*, Kerkwerve, The Netherlands: Backhuys Publishers.
- Finarelli, J.A. & Flynn, J.J., 2007. The evolution of encephalization in caniform carnivorans. *Evolution*, 61(7), pp.1758–1772.
- Finnegan, S. et al., 2011. The magnitude and duration of Late Ordovician–Early Silurian glaciation. *Science*, 311(6019), pp.903–906.
- Fitzgerald, P.C. & Carlson, S.J., 2006. Examining the latitudinal diversity gradient in Paleozoic terebratulide brachiopods: should singleton data be removed? *Paleobiology*, 32(3), pp.367–386.
- Flury, B.A., 1993. Estimation of principal points. *Applied Statistics*, 42, pp.139–151.
- Flury, B.A., 1990. Principal points. *Biometrika*, 77, pp.33–41.
- Foote, M., 1993a. Contributions of individual taxa to overall morphological disparity. *Paleobiology*, 19(4), pp.403–419.
- Foote, M., 1993b. Discordance and concordance between morphological and taxonomic diversity. *Paleobiology*, 19(2).
- Foote, M., 1996. Models of morphological diversification. In D. Jablonski, D. H. Erwin, & J. H. Lipps, eds. *Evolutionary Paleobiology*. University of Chicago Press, pp. 62–89.

- Foote, M., 1991a. Morphological and taxonomic diversity in clade's history: the blastoid record and stochastic simulations. *Contributions from the Museum of Paleontology*, 28(6), pp.101–140.
- Foote, M., 1994a. Morphological disparity in Ordovician-Devonian crinoids and the early saturation of morphological space. *Paleobiology*, 20(3).
- Foote, M., 1995. Morphological diversification of Paleozoic crinoids. *Paleobiology*, 21(3), pp.273–299.
- Foote, M., 1999. Morphological diversity in the evolutionary radiation of Paleozoic and post-Paleozoic crinoids. *Paleobiology*, 25(2).
- Foote, M., 1991b. Morphological patterns of diversification - examples from trilobites. *Palaeontology*, 34, pp.461–485.
- Foote, M., 1994b. Morphology of Ordovician-Devonian crinoids. *Contributions from the Museum of Paleontology University of Michigan*, 29(1), pp.1–39.
- Foote, M., 1990. Nearest-neighbor analysis of trilobite morphospace. *Systematic Zoology*, 39(4).
- Foote, M., 2000. Origination and extinction components of taxonomic diversity: Paleozoic and post-Paleozoic dynamics. *Paleobiology*, 26(4), pp.578–605.
- Foote, M., 1992a. Paleozoic record of morphological diversity in blastozoan echinoderms. *Proceedings of the National Academy of Sciences of the United States of America*, 89(16).
- Foote, M., 1992b. Rarefaction analysis of morphological and taxonomic diversity. *Paleobiology*, 18(1).
- Foote, M., 1997. The evolution of morphological diversity. *Annual Review of Ecology and Systematics*, 28(1), pp.129–152.
- Forrest, R., 2011. The plesiosaur site. Available at: <http://www.plesiosaur.com/> [Accessed February 18, 2011].
- Fortelius, M., 2011. Neogene of the old world database of fossil mammals (NOW). Available at: [www.helsinki.fi/science/now/](http://www.helsinki.fi/science/now/) [Accessed November 27, 2011].
- Fortey, R.A., Briggs, D.E.G. & Wills, M.A., 1996. The Cambrian evolutionary “explosion”: decoupling cladogenesis from morphological disparity. *Biological Journal of the Linnean Society*, 57(1), pp.13–33.
- Fortey, R.A. & Chatterton, B.D.E., 1988. Classification of the trilobite suborder Asaphina. *Paleontology*, 31, pp.165–222.
- Foth, C., Brusatte, S.L. & Butler, R.J., 2012. Do different disparity proxies converge on a common signal? Insights from the cranial morphometrics and evolutionary history of Pterosauria (Diapsida: Archosauria). *Journal of evolutionary biology*, 25(5), pp.904–915.
- Fraiser, M.L. & Bottjer, D.J., 2007. When bivalves took over the world. *Paleobiology*, 33(3), pp.397–413.

- Friedman, J.H., 1991. Multivariate adaptive regression splines. *Annals of Statistics*, 19, pp.1–67.
- Friedman, M., 2010. Explosive morphological diversification of spiny-finned teleost fishes in the aftermath of the end-Cretaceous extinction. *Proceedings of the Royal Society B: Biological Sciences*.
- Frobisch, J., 2008. Global taxonomic diversity of anamniotes (Tetrapoda, Therapsida) and the terrestrial rock record across the Permian-Triassic boundary. *PLoS ONE*, 3(11), pp.1–14.
- Galis, F., 1999. Why do almost all mammals have seven cervical vertebrae? Developmental constraints, hoxgenes, and cancer. *Journal of Experimental Zoology*, 285(1), pp.19–26.
- Gaston, K.J., 2000. Global patterns in biodiversity. *Nature*, 405(6783), pp.220–227.
- Gavrillets, S., 1999. Dynamics of clade diversification on the morphological hypercube. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 266(1421), pp.817–824.
- Gavrillets, S. & Losos, J.B., 2009. Adaptive radiation: contrasting theory with data. *Science*, 323(5915), pp.732–737.
- Geary, D.H. et al., 2010. The paradox of gradualism: phyletic evolution in two lineages of lymnocyprid bivalves (Lake Pannon, central Europe). *Paleobiology*, 36(4), pp.592–614.
- Gentry, A.H., 1988. Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Annals of the Missouri Botanical Garden*, 75(1), pp.1–34.
- Gerber, S., Eble, G.J. & Neige, P., 2008. Allometric space and allometric disparity: a developmental perspective in the macroevolutionary analysis of morphological disparity. *Evolution*, 62(6), pp.1450–1457.
- Ghiselin, M.T., 2005. The Darwinian revolution as viewed by a philosophical biologist. *Journal of Historical Biology*, 38(1), pp.123–136.
- Girard, C. & Renaud, S., 2012. Disparity changes in 370 Ma Devonian fossils: The signature of ecological dynamics? . *PLoS ONE*, 7, pp.1–12.
- Giribet, G. & Wheeler, W., 2002. On bivalve phylogeny: a high-level analysis of the Bivalvia (Mollusca) based on combined morphology and DNA sequence data. *Invertebrate Biology*, 121(4), pp.271–324.
- Good, I.J., 1953. The population frequencies of species and the estimation of population parameters. *Biometrika*, 40(3/4), pp.237–264.
- Gould, S.J., 1993. How to analyze Burgess shale disparity-a reply to Ridley. *Paleobiology*, 19(4). Available at: <http://dx.doi.org/10.2307/2401070>.
- Gould, S.J., 1991. The disparity of the burgess shale arthropod fauna and the limits of cladistic analysis: why we must strive to quantify morphospace. *Paleobiology*, 17(4).
- Gould, S.J. et al., 1977. The shape of evolution: a comparison of real and random clades. *Paleobiology*, 3(1), pp.23–40.

- Gould, S.J., 1989. *Wonderful life: the Burgess shale and the nature of history* W. W. Norton & C. N. York., eds., New York: W. W. Norton & Co.
- Gould, S.J. & Calloway, C.B., 1980. Clams and brachiopods - ships that pass in the night. *Paleobiology*, 6(4), pp.383–396.
- Gould, S.J., Gilinsky, N.L. & German, R.Z., 1987. Asymmetry of lineages and the direction of evolutionary time. *Science*, 236(4807), pp.1437–1441.
- Gradstein, F.M., Ogg, J.G. & Smith, A.G., 2004. A geologic time scale 2004.
- Grey, M., Haggart, J.W. & Smith, P.L., 2010. Morphological variability in time and space: an example of patterns within buchiid bivalves (Bivalvia, Buchiidae). *Palaeontology*, pp.1269–1280.
- Hallam, A., 1986. The Pliensbachian and Tithonian extinction events. *Nature*, 319(6056), pp.765–768.
- Hallam, A. & Wignall, P.B., 1997. *Mass extinctions and their aftermath*, Oxford: Oxford University Press.
- Hammer, Ø. & Harper, D.A.T., 2006. *Paleontological data analysis*, Blackwell Publishers Ltd.
- Hannisdal, B. & Peters, S.E., 2011. Phanerozoic Earth system evolution and marine biodiversity. *Science*, 334(6059), pp.1121–1124.
- Haq, B.U. & Schutter, S.R., 2002. A chronology of Paleozoic sea-level changes. *Science*, 322(5898), pp.64–68.
- Harmon, L.J. et al., 2010. Early bursts of body size and shape evolution are rare in comparative data. *Evolution*, 64(8), pp.2385–2396.
- Harmon, L.J. et al., 2003. Tempo and mode of evolutionary radiation in iguanian lizards. *Science*, 301(5635), pp.961–964.
- Harper, E.M., Taylor, J.D. & Crame, J.A., 2000. *Evolutionary biology of the Bivalvia*, London: The Geological Society.
- Hay, W.W. & Floegel, S., 2012. New thoughts about the Cretaceous climate and oceans. *Earth-science Reviews*, 15(4), pp.262–272.
- Hillebrand, H., 2004. On the generality of the latitudinal diversity gradient. *The American Naturalist*, 163(2), pp.192–211.
- Hoffmann, R., 2010. New insights on the phylogeny of the Lytoceratoidea (Ammonitina) from the septal lobe and its functional interpretation. *Revue de Paléobiologie*, Genève, 1, pp.1–156.
- Hone, D.W.E. & Benton, M.J., 2005. The evolution of large size: how does Cope's Rule work? *Trends in Ecology & Evolution*, 20(1), pp.4–6.
- Hopkins, M.J. & Lidgard, S., 2012. Evolutionary mode routinely varies among morphological traits within fossil species lineages. *Proceedings of the National Academy of Sciences*, 109(50), pp.20520–20525.

- Hu, D. et al., 2009. A pre-Archaeopteryx troodontid theropod from China with long feathers on the metatarsus. *Nature*, 461(7264), pp.640–643.
- Hubbard, R., 1978. The probable consequences of violating the normality assumption in parametric statistical analysis. *The Royal Geographical Society*, 10(5), pp.393–398.
- Huber, M., 2010. *Compendium of bivalves. A full-color guide to 3,300 of the World's Marine Bivalves. A status on Bivalvia after 250 years of research*, ConchBooks.
- Hughes, M., Gerber, S. & Wills, M.A., 2013. Clades reach highest morphological disparity early in their evolution. *Proceedings of the National Academy of Sciences*.
- Hunt, G., 2008. Evolutionary patterns within fossil lineages: model-based assessment of modes, rates, punctuations and process. In R. K. Bambach & P. H. Kelley, eds. *From Evolution to Geobiology: Research Questions Driving Paleontology at the Start of a New Century*. pp. 578–601.
- Hunt, G., 2006. Fitting and comparing models of phyletic evolution: random walks and beyond. *Paleobiology*, 32(4), pp.578–601.
- Hunt, G., Bell, M.A. & Travis, M.P., 2008. Evolution towards a new adaptive optimum: Phenotypic evolution in a fossil stickleback lineage. *Evolution*, 62(3), pp.700–710.
- Huntley, J., Xiao, S. & Kowalewski, M., 2006. 1.3 Billion years of acritarch history: An empirical morphospace approach. *Precambrian Research*, 144(1-2), pp.52–68.
- Isozaki, Y. & Aljinović, D., 2009. End-Guadalupian extinction of the Permian gigantic bivalve Alatoconchidae: End of gigantism in tropical seas by cooling. *Palaeogeography Palaeoclimatology Palaeoecology*, 284(1-2), pp.11–21.
- Van Iten, H. et al., 2013. Alternative interpretations of some earliest Ediacaran fossils from China. *Acta Palaeontologica Polonica*, 58(1), pp.111–113.
- Jablonski, D., 2005. Mass extinctions and macroevolution. *Paleobiology*, 31(2), pp.192–210.
- Jablonski, D. et al., 2013. Out of the tropics, but how? Fossils, bridge species, and thermal ranges in the dynamics of the marine latitudinal diversity gradient. *Proceedings of the National Academy of Sciences*, 110(26), pp.10487–10494.
- Jablonski, D. et al., 2003. The impact of the pull of the Recent on the history of marine diversity. *Science*, 300(5622), pp.1133–1135.
- Jablonski, D., Lidgard, S. & Taylor, P.D., 1997. Comparative ecology of bryozoan radiations: origin of novelties in cyclostomes and cheilostomes. *PALAIOS*, 12(6), pp.505–523.
- Jablonski, D., Roy, K. & Valentine, J.W., 2006. Out of the tropics: evolutionary dynamics of the latitudinal diversity gradient. *Science*, 314, pp.102–106.
- Jernvall, J., Hunter, J.P. & Fortelius, M., 1996. Molar tooth diversity, disparity, and ecology in Cenozoic ungulate radiations. *Science*, 274(5292), pp.1489–1492.
- Johnson, J.B. & Omland, K.S., 2004. Model selection in ecology and evolution. *Trends in Ecology & Evolution*, 19, pp.101–108.



- Johnston, P.A. & Goodbody, Q.H., 1989. Middle Devonian bivalves from Melville Island, Arctic Canada. *Canadian Society of Petroleum Geologists Memoir*, 14(3), pp.337–345.
- Kauffman, S., 1985. New questions in genetics and evolution. *Cladistics*, 1(3), pp.247–265.
- Kidwell, S.M., 2005. Shell composition has no net impact on large-scale evolutionary patterns in mollusks. *Science*, 307(5711), pp.914–917.
- Kiessling, W. et al., 2013. Taxonomic occurrences of Bivalvia. *Paleobiology Database*. Available at: <http://paleodb.org> [Accessed May 28, 2013].
- Kiessling, W. et al., 2012. Taxonomic occurrences of Bivalvia. *Paleobiology Database*. Available at: <http://paleodb.org> [Accessed March 13, 2012].
- Kiessling, W., Simpson, C. & Foote, M., 2010. Reefs as cradles of evolution and sources of biodiversity in the Phanerozoic. *Science*, 327(5962), pp.196–198.
- Kitchell, J.A. & MacLeod, N., 1989. Asymmetries of clade shape and the direction of evolutionary time: response. *Science*, 243(4898).
- Kitchell, J.A. & MacLeod, N., 1988. Macroevolutionary interpretations of symmetry and synchronicity in the fossil record. *Science*, 240(4856).
- Klingenberg, C.P. & Froese, R., 1991. A multivariate comparison of allometric growth-patterns. *Systematic Zoology*, 40, pp.410–419.
- Knoll, A.H. & Carroll, S.B., 1999. Early animal evolution: emerging views from comparative biology and geology. *Science*, 284(5423), pp.2129–2137.
- Kocot, K.M. et al., 2011. Phylogenomics reveals deep molluscan relationships. *Nature*, 477(7365), pp.542–546.
- Kolbe, S.E., Lockwood, R. & Hunt, G., 2011. Does morphological variation buffer against extinction? A test using veneroid bivalves from the Plio-Pleistocene of Florida. *Paleobiology*, 37(3), pp.355–368.
- Kosnik, M.A. et al., 2011. Changes in shell durability of common marine taxa through the Phanerozoic: evidence for biological rather than taphonomic drivers. *Paleobiology*, 37(2), pp.303–331.
- Kriz, J., 2004. Latest Frasnian and earliest Famennian (Late Devonian) bivalves from the Montagne Noire (France). *Senckenbergiana lethaea*, 84(1-2), pp.85–123.
- Kroh, A. & Mooi, R., 2011. World Echinoidea database version 2.0. Available at: [www.echinoidb.org](http://www.echinoidb.org) [Accessed May 19, 2011].
- Krug, A.Z. et al., 2009. Generation of Earth's first-order biodiversity pattern. *Astrobiology*, 9(1), pp.113–124.
- Krug, A.Z., Jablonski, D. & Valentine, J.W., 2007. Contrarian clade confirms the ubiquity of spatial origination patterns in the production of latitudinal diversity gradients. *Proceedings of the National Academy of Sciences of the United States of America*, 104(46), pp.18129–34.

- Kullman, P., 2011. GONIAT online Paleozoic ammonoid database system. Available at: [www.goniat.org/](http://www.goniat.org/) [Accessed June 16, 2011].
- Labandeira, C.C. & Eble, G.J., 2007. The fossil record of insect diversity and disparity. In J. Anderson et al., eds. *Gondwana alive: biodiversity and the evolving biosphere*. Witwatersrand University Press, Johannesburg.
- Labandeira, C.C. & Sepkoski, J.J., 1993. Insect diversity in the fossil record. *Science*, 261(5119), pp.310–315.
- Lamsdell, J.C., Braddy, S.J. & Tetlie, O.E., 2010. The systematics and phylogeny of the Stylonurina (Arthropoda: Chelicerata: Eurypterida). *Journal of Systematic Palaeontology*, 8(1), pp.49–61.
- Lamsdell, J.C., Percival, I.G. & Poschmann, M., 2013. The problematic “chelicerate” *Melbournopterus crossotus* Caster & Kjellesvig-Waering: a case of mistaken identity. *Alcheringa*, 37(3), pp.344–348.
- Leckie, R.M., Bralower, T.J. & Cashman, R., 2002. Oceanic anoxic events and plankton evolution: Biotic response to tectonic forcing during the mid-Cretaceous. *Palaeogeography*, 17(3), pp.1–29.
- Lefebvre, B. et al., 2006. Diversification of atypical Paleozoic echinoderms: a quantitative survey of patterns of stylophoran disparity, diversity, and geography. *Paleobiology*, 32(3), pp.483–510.
- Levinton, J. & Mackie, J., 2013. Latitudinal diversity relationships of fiddler crabs: biogeographic differences united by temperature. *Global Ecology and Biogeography*, 22(9), pp.1050–1059.
- Liljedahl, L., 1990. *Evyana*, new name for the Silurian bivalve *Fylgia* Liljedahl, 1989, non Kirby, 1889. *Geologiska Föreningen i Stockholm Förhandlingar*, 112(2), p.168.
- Liu, C., 1995. Jurassic bivalve palaeobiogeography of the proto-Atlantic and the application of multivariate analysis methods to palaeobiogeography. *Beringeria*, 16, pp.3–123.
- Lloyd, G.T., 2011. A refined modelling approach to assess the influence of sampling on palaeobiodiversity curves: new support for declining Cretaceous dinosaur richness. *Biology Letters*.
- Lloyd, G.T. et al., 2012. Sampling bias and the fossil record of planktonic foraminifera on land and in the deep sea. *Paleobiology*, 38(4), pp.569–584.
- Lloyd, G.T. & Friedman, M., 2013. A survey of palaeontological sampling biases in fishes based on the Phanerozoic record of Great Britain. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 372, pp.5–17.
- Lloyd, G.T., Smith, A.B. & Young, J.R., 2011. Quantifying the deep-sea rock and fossil record bias using coccolithophores. In *Comparing the Geological and Fossil Records: Implications for Biodiversity Studies*. Geological Society, London, Special Publications, pp. 167–177.
- Lomolino, M. V. et al., 2013. Of mice and mammoths: generality and antiquity of the island rule. *Journal of Biogeography*, 40(8), pp.1427–1439.

- Lupia, R., 1999. Discordant morphological disparity and taxonomic diversity during the Cretaceous angiosperm radiation; North American pollen record. *Paleobiology*, 25(1), pp.1–28.
- Mah, C.L., 2009. World Asteroidea database. Available at: <http://www.marinespecies.org/asteroidea> [Accessed April 20, 2011].
- Mannion, P.D. et al., 2011. Testing the effect of the rock record on diversity: a multidisciplinary approach to elucidating the generic richness of sauropodomorph dinosaurs through time. *Biological Reviews*, 86(1), pp.157–181.
- Markov, A. V., Bondarev, A.A. & Vinarsky, M. V., 2012. Diversity dynamics of echinoderms and evolution of marine communities. *Paleontological Journal*, 46(8), pp.865–876.
- Markwick, P.J., 2002. Integrating the present and past records of climate, biodiversity and biogeography: implications for palaeoecology and palaeoclimatology. In J. A. Crame & A. W. Owen, eds. *Palaeobiogeography and Biodiversity Change: the Ordovician and Mesozoic-Cenozoic Radiations*. London: The Geological Society, pp. 179–199.
- Marshall, C.R. & Valentine, J.W., 2010. The importance of preadapted genomes in the origin of the animal body plans and the Cambrian explosion. *Evolution*, 64(5), pp.1189–1201.
- Mayhew, P.J. et al., 2012. Biodiversity tracks temperature over time. *Proceedings of the National Academy of Sciences of the United States of America*, 109(38), pp.15141–15145.
- McShea, D.W., 1996. Perspective: metazoan complexity and evolution: is there a trend? *Evolution*, 50(2), pp.477–492.
- McShea, D.W., 1998. Possible largest-scale trends in organismal evolution: eight “live hypotheses.” *Annual Review of Ecology and Systematics*, 29, pp.293–318.
- McShea, D.W. & Brandon, R.N., 2010. *Biology's first law: the tendency for diversity and complexity to increase in evolutionary systems*, The University of Chicago Press.
- Mikkelsen, P.M., 2011. Speciation in modern marine bivalves (Mollusca: Bivalvia): insights from the published record. *American Malacological Bulletin*, 29(1/2), pp.217–245.
- Miller, A.I., 1989. Spatio-temporal transitions in Paleozoic Bivalvia: A field comparison of upper Ordovician and upper Paleozoic bivalve-dominated fossil assemblages. *Historical Biology: An International Journal of Paleobiology*, 2(3), pp.227–260.
- Miller, A.I. & Foote, M., 1996. Calibrating the Ordovician radiation of marine life; implications for Phanerozoic diversity trends. *Paleobiology*, 22(2), pp.304–309.
- Miller, A.I. & Sepkoski, J.J., 1988. Modeling bivalve diversification: the effect of interaction on a macroevolutionary system. *Paleobiology*, 14(4), pp.364–369.
- Mitchell, C.E., 1987. Evolution and phylogenetic classification of the Diplograptacea. *Palaeontology*, 30(2), pp.353–405.
- Mittelbach, G.G. et al., 2007. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecology Letters*, 10(4), pp.315–331.

- Moreau, C.S. & Bell, C.D., 2013. Testing the museum versus cradle tropical biological diversity hypothesis: phylogeny, diversification, and ancestral biogeographic range evolution of the ants. *Evolution*, 67(8), pp.2240–2257.
- Munnecke, A. et al., 2010. Ordovician and Silurian sea–water chemistry, sea level, and climate: A synopsis. *Palaeogeography Palaeoclimatology Palaeoecology*, 296(3-4), pp.389–413.
- Mutterlose, J., Bornemann, A. & Herrle, J., 2009. The Aptian – Albian cold snap: evidence for “mid” Cretaceous icehouse interludes. *Neues Jahrbuch Fur Geologie Und Palaontologie-Abhandlungen*, 252(2), pp.217–225.
- Naimark, E.B. & Markov, A. V., 2011. Northward shift in faunal diversity: a general pattern of evolution of Phanerozoic marine biota. *Biological Bulletin Reviews*, 1(1), pp.71–81.
- Navarro, J.M. et al., 2008. Feeding behavior of the infaunal bivalve *Tagelus dombeii* (Lamarck, 1818). Suspension vs. deposit feeding. *Revista de Biología Marina y Oceanografía*, 43(3), pp.599–605.
- Neige, P., 2003. Spatial patterns of disparity and diversity of the Recent cuttlefishes (Cephalopoda) across the Old World. *Journal of Biogeography*, 30(8), pp.1125–1137.
- Neustupa, J., Cerna, K. & St’astny, J., 2009. Diversity and morphological disparity of desmid assemblages in Central European peatlands. *Hydrobiologia*, 630(1), pp.243–256.
- Neveskaya, L.A., 2008. Dynamics of taxonomic diversity of bivalves in the Phanerozoic. *Paleontological Journal*, 42(4), pp.335–342.
- Norris, R.D. et al., 2002. Jiggling the tropical thermostat in the Cretaceous hothouse. *Geology*, 30(4), pp.299–302.
- Novas, F.E. et al., 2009. A bizarre Cretaceous theropod dinosaur from Patagonia and the evolution of Gondwanan dromaeosaurids. *Proceedings of the Royal Society B: Biological Sciences*, 276(1659), pp.1101–1107.
- Nurnberg, S. & Aberhan, M., 2013. Habitat breadth and geographic range predict diversity dynamics in marine Mesozoic bivalves. *Paleobiology*, 39(3), pp.360–372.
- O’Connor, A., Moncrieff, C. & Wills, M.A., 2011. Variation in stratigraphic congruence (GER) through the Phanerozoic and across higher taxa is partially determined by sources of bias. In A. B. Smith & A. J. McGowan, eds. *Comparing the Geological and Fossil Records: Implications for Biodiversity Studies*. Geological Society Special Publication, pp. 31–52.
- Ockelmann, K.W. & Dinesen, G.E., 2011. Life on wood –the carnivorous deep-sea mussel *Idas argenteus* (Bathymodiolinae, Mytilidae, Bivalvia). *Marine Biology Research*, 7(1), pp.71–84.
- Ogg, J.G., Ogg, G. & Gradstein, F.M., 2008. The concise geologic time scale. In N. Woodcock, ed. *Geological Magazine*. Cambridge, New York, Melbourne: Cambridge University Press, pp. 156–177.
- Parkhaev, P.Y., 2006. The Early Cambrian radiation of Mollusca. In W. F. Ponder & D. R. Lindberg, eds. *Phylogeny and Evolution of the Mollusca*. Berkeley and Los Angeles, California: University of California Press, pp. 33–69.

- Peck, L.S., Powell, D.K. & Tyler, P.A., 2007. Very slow development in two Antarctic bivalve molluscs, the infaunal clam *Laternula elliptica* and the scallop *Adamussium colbecki*. *Marine Biology*, 150(6), pp.1191–1197.
- Perry, H. & Larsen, K., 2004. Guide to shelf invertebrates, Gulf of Mexico. Available at: <http://www.gsmfc.org/> [Accessed September 19, 2011].
- Peters, S.E., 2008. Environmental determinants of extinction selectivity in the fossil record. *Nature*, 454(7204), pp.626–638.
- Peters, S.E. & Foote, M., 2001. Biodiversity in the Phanerozoic: a reinterpretation. *Paleobiology*, 27(4), pp.583–601.
- Plazzi, F. et al., 2011. A molecular phylogeny of bivalve mollusks: ancient radiations and divergences as revealed by mitochondrial genes. *PLoS ONE*, 6(11), p.e27147.
- Polechova, M., 2013. Bivalves from the Middle Ordovician Sarka Formation (Prague Basin, Czech Republic). *Bulletin of Geosciences*, 88(2), pp.427–461.
- Pollitt, J.R., Fortey, R.A. & Wills, M.A., 2005. Systematics of the trilobite families Lichidae Hawle & Corda, 1847 and Lichakephalidae Tripp, 1957: The application of bayesian inference to morphological data. *Journal of Systematic Palaeontology*, 3(3), pp.225–241.
- Ponder, W. & Lindberg, D.R., 2006. Molluscan evolution and phylogeny an introduction. In W. Ponder & D. R. Lindberg, eds. *Phylogeny and Evolution of the Mollusca*. Berkeley and Los Angeles, California: University of California Press, pp. 1–19.
- Ponder, W. & Lindberg, D.R., 2008. *Phylogeny and evolution of the Mollusca*, California: University of California Press, USA.
- Prendini, L., Francke, O.F. & Vignoli, V., 2010. Troglomorphism, trichobothriotaxy and typhlochactid phylogeny (Scorpiones, Chactioidea): more evidence that troglobitism is not an evolutionary dead-end. *Cladistics*, 26(2), pp.117–142.
- Qian, H., Badgley, C. & Fox, D.L., 2009. The latitudinal gradient of beta diversity in relation to climate and topography for mammals in North America. *Global Ecology and Biogeography*, 18(1), pp.111–122.
- Raup, D.M. & Boyajian, G.E., 1988. Patterns of Generic Extinction in the Fossil Record. *Paleobiology*, 14(2), pp.109–125.
- Rode, A.L. & Lieberman, B.S., 2004. Using GIS to unlock the interactions between biogeography, environment, and evolution in Middle and Late Devonian brachiopods and bivalves. *Palaeogeography Palaeoclimatology Palaeoecology*, 211(3-4), pp.345–359.
- Ros, S. et al., 2011. Coping between crises: Early Triassic–early Jurassic bivalve diversity dynamics. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 311(3-4), pp.184–199.
- Ros, S. & Echevarria, J., 2011. Bivalves and evolutionary resilience: old skills and new strategies to recover from the P/T and T/J extinction events. *Historical Biology*, 23(4), pp.411–429.
- Ros, S. & De Renzi, M., 2005. Preservation biases, rates of evolution and coherence of databases: Bivalvia as a study case. *Ameghiniana*, 42(3), pp.549–558.

- Roy, K., Hunt, G., Jablonski, D., et al., 2009. A macroevolutionary perspective on species range limits. *Proceedings of the Royal Society B*, 276(1661), pp.1485–1493.
- Roy, K., Balch, D.P. & Hellberg, M.E., 2001. Spatial patterns of morphological diversity across the Indo-Pacific: analyses using strombid gastropods. *Proceedings: Biological Sciences*, 268(1485), pp.2503–2508.
- Roy, K. & Goldberg, E.E., 2007. Origination, extinction, and dispersal: integrative models for understanding present day diversity gradients. *The American Naturalist*, 170(2), pp.71–85.
- Roy, K., Hunt, G. & Jablonski, D., 2009. Phylogenetic conservatism of extinctions in marine bivalves. *Science*, 325(5941), pp.733–737.
- Roy, K., Jablonski, D. & Valentine, J.W., 2004. Beyond species richness: biogeography patterns and biodiversity dynamics using other metrics of diversity. In M. V. Lomolino & L. R. Heaney, eds. *Frontiers of Biogeography: New Directions in the Geography of Nature*. Sunderland: Sinauer, pp. 151–170.
- Roy, K., Jablonski, D. & Valentine, J.W., 2000. Dissecting latitudinal diversity gradients: functional groups and clades of marine bivalves. *Proceedings of the Royal Society B: Biological Sciences*, 267, pp.293–299.
- Royall, R.M., 1997. *Statistical evidence: a likelihood paradigm*, Chapman & Hall, New York.
- Ruban, D.A., 2006. Taxonomic diversity dynamics of the Jurassic bivalves in the Caucasus: Regional trends and recognition of global patterns. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 239(1-2), pp.63–74.
- Ruta, M., 2009. Patterns of morphological evolution in major groups of Palaeozoic Temnospondyli (Amphibia: Tetrapoda) M Ruta et al., eds. *Patterns and Processes in Early Vertebrate Evolution*, (81), pp.91–120.
- Ruta, M., Wagner, P.J. & Coates, M.I., 2006. Evolutionary patterns in early tetrapods. I. Rapid initial diversification followed by decrease in rates of character change. *Proceedings. Biological sciences / The Royal Society*, 273(1598), pp.2107–2111.
- Sahney, S. & Benton, M.J., 2008. Recovery from the most profound mass extinction of all time. *Proceedings of the Royal Society B*, 275(1636), pp.759–765.
- Sanchez, T.M., 2008. The early bivalve radiation in the Ordovician Gondwanan basins of Argentina. *Alcheringa*, 32(3), pp.223–246.
- Sander, P.M. et al., 2011. Biology of the sauropod dinosaurs: the evolution of gigantism. *Biological Reviews*, 86(1), pp.117–155.
- Sansom, R.S., Gabbott, S.E. & Purnell, M.A., 2011. Decay of vertebrate characters in hagfish and lamprey (Cyclostomata) and the implications for the vertebrate fossil record. *Proceedings of the Royal Society B Biological Sciences*, 278(1709), pp.1150–1157.
- Sartori, A.F. & Harper, E.M., 2009. Sticky bivalves from the Mesozoic: clues to the origin of the anomalodesmatan arenophilic system. *Lethaia*, 42, pp.486–494.
- Saunders, W.B. & Work, D.M., 1996. Shell morphology and suture complexity in Upper Carboniferous ammonoids. *Paleobiology*, 22(2), pp.189–218.

- Schemske, D.W., 2009. Biotic interactions and speciation in the tropics. In R. K. Butlin, J. R. Bridle, & D. Schuller, eds. *Speciation and Patterns of Diversity*. Cambridge University Press, pp. 219–239.
- Schmidt, R., 2011. Recent and fossil bryozoa. *International Bryozoology Association*. Available at: [www.bryozoa](http://www.bryozoa). [Accessed June 15, 2011].
- Schneider, J.A., 1999. A new species of cardiid bivalve from the Late Cretaceous of Jamaica. *Caribbean Journal of Science*, 35(3-4), pp.221–224.
- Schneider, S. et al., 2010. Ecophenotypic plasticity versus evolutionary trends - morphological variability in Upper Jurassic bivalve shells from Portugal. *Acta Palaeontologica Polonica*, 55(4), pp.701–732.
- Sepkoski, J.J., 2002. *A compendium of fossil marine genera* D. Jablonski & M. Foote, eds., Paleontological Research Institution.
- Sepkoski, J.J.J., 1984. A kinetic model of Phanerozoic taxonomic diversity. III. Post-Paleozoic families and mass extinctions. *Paleobiology*, 10(2).
- Sharma, P.P. et al., 2013. Into the deep: A phylogenetic approach to the bivalve subclass Protobranchia. *Molecular Phylogenetics and Evolution*, 69(1), pp.188–204.
- Sheehan, P.M., 1977. Species diversity in the Phanerozoic: a reflection of labor by systematists. *Paleobiology*, 3, pp.325–328.
- Shen, B. et al., 2008. The Avalon explosion: evolution of Ediacara morphospace. *Science*, 319(5859), pp.81–84.
- Sidlauskas, B., 2008. Continuous and arrested morphological diversification in sister clades of characiform fishes: a phylomorphospace approach. *Evolution*, 62(12), pp.3135–3156.
- Simpson, G.G., 1944. *Tempo and mode in evolution*, New York: Columbia University Press.
- Sites Jr., J.W. et al., 1996. Character Congruence and Phylogenetic Signal in Molecular and Morphological Data Sets: A Case Study in the Living Iguanas (Squamata, Iguanidae). *Molecular Biology and Evolution*, 13(8), pp.1087–1105.
- Slater, G.J. et al., 2010. Diversity versus disparity and the radiation of modern cetaceans. *Proceedings of the Royal Society B: Biological Sciences*, 277(1697), pp.3097–3104.
- Slieker, F., 2011. Natural History Museum Rotterdam. Available at: <http://www.nmr-pics.nl/> [Accessed April 13, 2011].
- Smith, A.B., 1994. *Systematics and the Fossil Record*, Oxford: Blackwell Science.
- Smith, A.B. & Kroh, A., 2011. The echinoid directory. Available at: <http://www.nhm.ac.uk/research-curation/research/projects/echinoid-directory/index.html> [Accessed May 19, 2011].
- Smith, A.B. & McGowan, A.J., 2007. The shape of the Phanerozoic marine palaeodiversity curve: how much can be predicted from the sedimentary rock record of western Europe? *Palaeontology*, 50, pp.765–774.

- Smith, A.B. & McGowan, A.J., 2011. The ties linking rock and fossil records and why they are important for palaeobiodiversity studies. In A. B. Smith & A. J. McGowan, eds. *Comparing the Geological and Fossil Records: Implications for Biodiversity Studies*. Geological Society, London, Special Publications, pp. 1–7.
- Smith, A.S., 2011. The Plesiosaur Directory. Available at: [www.plesiosauria.com/](http://www.plesiosauria.com/) [Accessed February 18, 2011].
- Smith, J.M. et al., 1985. Developmental constraints and evolution: a perspective from the mountain lake conference on development and evolution. *The Quarterly Review of Biology*, 60(3), pp.265–287.
- Smith, K.L. et al., 2011. Evidence of constrained phenotypic evolution in a cryptic species complex of agamid lizards. *Evolution*, 65(4), pp.976–992.
- Snedden, J.W. & Liu, C., 2010. A compilation of Phanerozoic sea-level change, coastal onlaps and recommended sequence designations. *AAPG Search and Discovery*, 40594, pp.1–3.
- Sousa, R., Gutierrez, J.L. & Aldridge, D.C., 2009. Non-indigenous invasive bivalves as ecosystem engineers. *Biological Invasions*, 11, pp.2367–2385.
- Sprinkle, J. & Kier, P.M., 1987. Phylum Echinodermata - introduction to Blastoidea. In R. S. Boardman, A. H. Cheetham, & A. J. Rowell, eds. *Fossil Invertebrates*. Palo Alto: Blackwell Scientific.
- Stanley, S.M., 1968. Post-Paleozoic adaptive radiation of infaunal bivalve molluscs - a consequence of mantle fusion and siphon formation. *Journal of Paleontology*, 42(1), pp.214–229.
- Stanley, S.M. & Yang, X., 1987. Approximate evolutionary stasis for bivalve morphology over millions of years: a multivariate, multilineage study. *Paleobiology*, 13(2), pp.113–139.
- Stebbins, G.L., 1974. *Flowering plants: evolution above the species level*, Belknap Press of Harvard University Press.
- Swofford, D.L., 2003. PAUP\*. Phylogenetic analysis using parsimony (\*and other methods). Version 4.
- Team, R.D.C., 2012. R: A language and environment for statistical computing.
- Thomas, R.D.K., Shearman, R.M. & Stewart, G.W., 2000. Evolutionary exploitation of design options by the first animals with hard skeletons. *Science*, 288(5469), pp.1239–1242.
- Thorne, P.M., Ruta, M. & Benton, M.J., 2011. Resetting the evolution of marine reptiles at the Triassic-Jurassic boundary. *Proceedings of the National Academy of Sciences of the United States of America*, 108(20), pp.8339–8344.
- Ubukata, T., 2003. A theoretical morphologic analysis of bivalve ligaments. *Paleobiology*, 29(3), pp.369–380.
- Uetz, P., 2011. The reptile database. Available at: <http://www.reptile-database.org> [Accessed June 14, 2011].



- Uhen, M.D., 1996. An evaluation of clade-shape statistics using simulations and extinct families of mammals. *Paleobiology*, 22(1).
- Valentich-Scott, P. & Dinesen, G.E., 2004. Rock and coral boring Bivalvia (Mollusca) of the middle Florida keys, U.S.A. *Malacologia*, 46(2), pp.339–354.
- Valentine, J.W. et al., 2008. Incumbency, diversity, and latitudinal gradients. *Paleobiology*, 34(2), pp.169–178.
- Valentine, J.W. et al., 2013. The sampling and estimation of marine paleodiversity patterns: implications of a Pliocene model. *Paleobiology*, 39(1), pp.1–20.
- Valentine, J.W., 1995. Why no new phyla after the Cambrian? Genome and ecospace hypotheses revisited. *Palaos*, 10(2), pp.190–194.
- Valentine, J.W., Collins, A.G. & Meyer, C.P., 1994. Morphological complexity increase in metazoans. *Paleobiology*, 20(2), pp.131–142.
- Valentine, J.W., Erwin, D.H. & Jablonski, D., 1996. Developmental evolution of metazoan bodyplans: the fossil evidence. *Developmental Biology*, 173(2), pp.373–381.
- Valentine, J.W. & Jablonski, D., 2010. Origins of marine patterns of biodiversity: some correlates and applications. *Palaeontology*, 53(6), pp.1203–1210.
- Valentine, J.W., Jablonski, D. & Erwin, D.H., 1999. Fossils, molecules and embryos: new perspectives on the Cambrian explosion. *Development*, 126(5), pp.851–859.
- Vilhena, D.A. et al., 2013. Bivalve network reveals latitudinal selectivity gradient at the end-Cretaceous mass extinction. *Nature Scientific Reports*, 3(1790), pp.1–5.
- Vinarskia, M. V., Bondarev, A.A. & Markov, A. V., 2011. Mollusks in Phanerozoic marine communities: implications from the analysis of global paleontological databases. *Paleontological Journal*, 45(4), pp.358–369.
- Waggoner, B. & Collins, A.G., 2004. Reductio ad absurdum: testing the evolutionary relationships of Ediacaran and Paleozoic problematic fossils using molecular divergence dates. *Journal of Paleontology*, 78(1), pp.51–61.
- Wagner, A., 2011. Genotype networks shed light on evolutionary constraints. *Trends in Ecology & Evolution*, 26(11), pp.577–584.
- Wagner, P.J., 2000. Exhaustion of morphologic character states among fossil taxa. *Evolution*, 54(2), pp.365–386.
- Wagner, P.J., 2010. Paleontological perspectives on morphological evolution. In M. A. Bell et al., eds. *Evolution since Darwin: the first 150 years*. Sunderland: Sinauer, pp. 451–578.
- Wagner, P.J., 1997. Patterns of morphologic diversification among the Rostroconchia. *Paleobiology*, 23(1), pp.115–150.
- Wagner, P.J., 1995. Testing evolutionary constraint hypotheses with Early Paleozoic gastropods. *Paleobiology*, 21(3).
- Wagner, P.J., Ruta, M. & Coates, M.I., 2006. Evolutionary patterns in early tetrapods. II. Differing constraints on available character space among clades. *Proceedings of the Royal Society B-Biological Sciences*, 273(1598), pp.2113–2118.

- Wallace, A. & Farrow, M., 1999. The pattern of variation in centipede segment number as an example of developmental constraint in evolution. *Journal of Theoretical Biology*, 200(2), pp.183–191.
- Wang, J.G., Lim, E.T. & Venkateswarlu, R., 2004. Stereo head/face detection and tracking. *Icip: 2004 International Conference on Image Processing, Vols 1- 5*, pp.605–608.
- Wang, X., Tedford, R.H. & Taylor, B.E., 1999. Phylogenetic systematics of the Borophaginae (Carnivora: Canidae). *Bulletin of the American Museum of Natural History*, 243, pp.1–391.
- Watkins, R., 1997. Paleoecology of Silurian reef bivalves, Racine Formation, North America. *Lethaia*, 29, pp.171–180.
- Watson, S.-A. et al., 2012. Marine invertebrate skeleton size varies with latitude, temperature and carbonate saturation: implications for global change and ocean acidification. *Global Change Biology*, 18(10), pp.3026–3038.
- Webster, M., 2007. A Cambrian peak in morphological variation within trilobite species. *Science*, 317(5837), pp.499–502.
- Wesley-Hunt, G.D., 2005. The morphological diversification of carnivores in North America. *Paleobiology*, 31(1), pp.35–55.
- Wills, M.A., 1998a. Cambrian and Recent disparity: the picture from priapulids. *Paleobiology*, 24, pp.177–199.
- Wills, M.A., 1998b. Crustacean disparity through the Phanerozoic: comparing morphological and stratigraphic data. *Biological Journal of the Linnean Society*, 65(4), pp.455–500.
- Wills, M.A., 2001. Morphological disparity a primer. In J. M. Adrain, G. D. Edgecombe, & B. S. Lieberman, eds. *Fossils, Phylogeny and Form: an analytical approach*. New York: Kluwer Academic/Plenum Publishers, p. 402.
- Wills, M.A. et al., 2012. The disparity of priapulid, archaeopriapulid and palaeoscolecoid worms in the light of new data. *Journal of evolutionary biology*.
- Wills, M.A. et al., 1995. The significance of fossils in understanding arthropod evolution. *Verhandlungen der Deutschen Zoologischen Gesellschaft*, 88(2), pp.203–215.
- Wills, M.A., Briggs, D.E.G. & Fortey, R.A., 1994. Disparity as an evolutionary index: a comparison of Cambrian and Recent arthropods. *Paleobiology*, 20(2).
- Wills, M.A. & Fortey, R.A., 2000. The shape of life: how much is written in stone? *BioEssays*, 22(12), pp.1142–1152.
- Wilson, D.E. & Reeder, D.M., 2005. *Mammal Species of the World. A Taxonomic and Geographic Reference* 3rd ed., Johns Hopkins University Press.
- Wimsatt, W.A., 1986. *Integrating scientific disciplines* W. Bechtel, ed., New York: Springer.
- Zelditch, M.L., Sheets, H.D. & Fink, W.L., 2003. The ontogenetic dynamics of shape disparity. *Paleobiology*, 29(1), pp.139–156.

Zelditch, M.L. et al., 2004. *Geometric morphometrics for biologists: a primer*, San Diego: Elsevier Academic Press.

Zengqiong, H. & Gangsheng, Z., 2011. A new structural model of bivalve ligament from *Solen grandis*. *Micron*, 42(7), pp.706–711.

## ***Appendices***

On Disc:

Appendix 1 – Chapter 1 raw data, R code, Literature table 1, results table 2 and character state exploration table 3.

Appendix 2 – Chapter 3 R code, raw data and results, additional analyses, Mixed Range and Treatise stratigraphy table 1 and morphological treatments table 2

Appendix 3 – Chapter 4 R code, raw data and results